

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
19 August 2004 (19.08.2004)

PCT

(10) International Publication Number
WO 2004/070040 A1

- (51) International Patent Classification⁷: C12N 15/82, 15/90, 15/65
- (74) Agents: HELBING, Jörg et al.; Von Kreisler Selting
Werner, Postfach 10 22 41, 50462 Köln (DE).
- (21) International Application Number:
PCT/EP2004/000968
- (81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD,
MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG,
PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM,
ZW.
- (22) International Filing Date: 3 February 2004 (03.02.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
03002424.4 5 February 2003 (05.02.2003) EP
60/446,921 12 February 2003 (12.02.2003) US
- (84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), Euro-
pean (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR,
GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
ML, MR, NE, SN, TD, TG).
- (71) Applicant (for all designated States except US):
ARTEMIS PHARMACEUTICALS GMBH [DE/DE];
Neurather Ring 1, 51062 Köln (DE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): KAUSELMANN,
Gunter [DE/DE]; Waisenhausgasse 48, 50676 Köln
(DE). ZEVIK, Branko [DE/DE]; Gezelinallee 13,
51375 Leverkusen (DE). SEIBLER, Jost [DE/DE];
Leipziger Platz 13, 50733 Köln (DE). KERN, Heidrun
[DE/DE]; Sonnenstr. 21, 40227 Düsseldorf (DE).
- Published:
— with international search report
- For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: AUTOMATED GENE-TARGETING USING NON-TOXIC DETECTABLE MARKERS

(57) Abstract: The invention relates to a method that enables automated identification and isolation of cells harbouring a predetermined genetic modification (homologous recombination) using detectable/sortable markers, e.g. fluorescence markers, to identify homologous DNA modifications. Suitable vectors are also provided.

WO 2004/070040 A1

BEST AVAILABLE COPY

Automated Gene-Targeting Using Non-Toxic Detectable Markers

The invention relates to a method that enables automated identification and isolation of cells harbouring a predetermined genetic modification (homologous recombination) using detectable/sortable markers, e.g. fluorescence markers, to identify homologous DNA modifications. Suitable vectors are also provided.

Background

Gene targeting in cells, including embryonic stem (ES) cells, relies on the homologous recombination (HR) between a native chromosomal gene and introduced exogenous DNA (Smithies, O. et al., *Nature*, 317(6034):230-4 (1985); Thomas, K. R. et al., *Cell*, 44(3):419-28 (1986); Doetschman, T. et al., *Proc. Natl. Acad. Sci. USA* 85(22):8583-7 (1988); Thomas, K. R., Capecchi, M. R., *Cell*, 51(3):503-12 (1987)). The method requires transfection of the foreign DNA (targeting vector) into the ES cells, usually by electroporation. HR is a very inefficient process, working in the range of 1×10^{-5} – 1×10^{-6} events per electroporated ES cells (for review see Hooper, M.C., Harwood, *Embryonal Stem Cell: Introducing Planned Changes into the Animal Geneline*, New York (1993)). Most genes of interest are not directly selectable. The identification of homologous recombined (gene targeted) ES cells therefore require the application of selection mechanism.

Positive selection is usually achieved by incorporation of an antibiotic resistance gene into the targeting vector, located between at least 2 stretches of DNA homologous to the targeting region. Upon addition of antibiotic to the culture medium of electroporated ES cells, expression of the positive selection marker allows survival only of those cells which have stably introduced the electroporated DNA into their genome.

Additionally, negative selection strategies further enrich selective growth and identification of HR events. A negative selection marker is added at the 5' or 3' end of a homologous DNA stretch (Mansour, S. L. et al., *Nature*, 336:348-352 (1988);

U.S. Patents Nos. 5,487,992 and 5,464,764). Non-homologous DNA regions outside the homologous regions are not incorporated into the genome if HR has occurred. Current protocols (U.S. Patents 5,487,992 and 5,464,765) require the addition of a compound to the culture medium, to mediate selection against incorporation of the marker, effectively killing ES cells, which have undergone random recombination. Negative selection generally achieves enrichments of only 2-5fold (Sedivy, J. M., Dutriaux, A., Trends Genet., 15(3):88-90 (1999)).

Given optimal vector design to allow homologous recombination, the described selection strategies known in the art (e.g. the positive-negative selection strategies disclosed in U.S. Patents 5,487,992 and 5,464,765) allow identification of HR in about 5% of all analysed ES cells. Said selection strategies are certainly not optimal and it is therefore desirable to

- (a) optimize enrichment of homologous recombinants,
- (b) reduce variability in efficiency of the currently available methods for enrichment,
- (c) avoid unwanted potential toxicity of agents for negative or negative and positive selection agents. Gancyclovir, the most widely used agent for negative selection, inhibits DNA polymerase and therefore cellular proliferation (Matthews, T. and Boehme, R., Reviews of Infectious Diseases, 10(3):490-4 (1998)). Such inhibition may not be tolerable under GMP standards, in particular in potential applications of gene targeting for gene therapy in humans.

Moreover, the selection procedures known in the art requires ES cell culture over a period of 7 - 8 days, to allow selection to work, and clonal expansion of cells to an absolute number of ca. 2000, to enable molecular analysis and further expansion of the cell population. However, culture of ES cells requires adherent growth of the cells on defined substrate, usually mitotically inactivated „feeder cells“. This, in turn, necessitates laborious, manual "picking" of identified ES clones from their substrate to enable individual clonal growth in culture vessels for maintenance and molecular analysis. It was therefore also desirable to reduce the time period required to isolate potential HR ES clones as well as the manual labour, i.e. picking of ES cell clones, i.e. by increasing the automation of the gene targeting process employing current protocols.

Fluorescent probes on the other side are known as a powerful tool for identification of molecular events in single cells. The most widely known is the green fluorescent protein (GFP) from bioluminescent jellyfish *Aequorea victoria*. However, approximately 30 distinct fluorescent proteins have been discovered and cloned from a variety of species (Labas, Y. A. et al., PNAS, 99(7):4256-4261 (2002)), (Matz, M. V. et al., Nat. Biotechnol., 17(10):906-18 (1999)) for review see: (Zhang, J. et al., Nat. Rev. Mol. Cell. Biol., 3(12):906-18 (2002)).

For certain reasons such as the sensitivity of the cells to be transfected, optically detectable markers including fluorescent probes such as GFP and its variants were only seldom employed as detectable markers in gene targeting experiments. E. g. WO 02/06630 discloses that GFP might be utilized instead of a neomycin phosphotransferase gene in a gene targeting process similar to the basic positive selection protocol. Furthermore WO 03/0022725 discloses that fluorescent gene cassettes can be used as an alternative to negative selection markers in order to further enrich ES cell clones which underwent homologous recombination.

A key property of murine ES cells is that they can be maintained indefinitely *in vitro* if cultured in the presence of the cytokine leukemia inhibitory factor (LIF) (Smith, A. G. and Hooper, M., Dev. Biol. 121:1-9 (1987); Williams, R. L. et al. (1988)). ES cells retain the capacity to participate normally in embryogenesis and contribute to all tissues of the mouse embryo when introduced into host blastocysts, including the germ line (reviewed by Robertson, E.J., Oxford, UK, IRL Press: 71-112 (1987); Smith, A.G., Sem. Cell. Biol. 3:385-399 (1992)). However, retention of germ-line transmission competence is often elusive. It depends absolutely on adherence to a rigorous tissue culture regime, with avoidance of any untoward selective pressures. Furthermore, ES cells are only modestly able to amplify if plated as single cells to form clonal populations. The plating efficiency is reported to range from 3 to 5% (Reid, L. H. et al., Mol. Cell. Biol. 11: 2767-2777 (1991); Templeton, N. S., et al., Gene Therapy 4(7): 700-9 (1997)).

Altogether, low plating efficiency, differentiation accompanied by loss of germ line-competence and subsequent failure to produce genetically modified mice is still the key obstacle for the employment of automated sorting methodology for ES cells in gene targeting studies.

Methods such as Fluorescence Activated Cell Sorting have been described for sorting of human (Schuldiner, M.J. et al., Stem Cells 21(3):257-265 (2003)) and murine (Reddy, S. et al., Proc. Natl. Acad. Sci. USA 89:6721-6725 (1992); WO 03/002272) ES cells, however such cells were only bulk sorted as pools of marker expressing cells for subsequent analysis.

High frequency gene targeting and isolation of single cells with defined characteristics in an automated fashion as described in this application is of great advantage to improve effectiveness and precision of gene targeting experiments.

Summary of the Invention

It was now found that certain vectors for targeted homologous recombination which contain one or more expression cassettes coding for a detectable marker and being placed outside the targeting cassette (i.e. the region of homology of the vector to the genomic DNA) allow rapid and reliable distinction, preferably visual distinction, between targeted and non-targeted ES cells in an automated fashion. One example of such an optically detectable marker is ZsGreen (Clontech, Palo Alto, CA). A particular example is the application of a first fluorescence marker gene (e.g. ZsGreen) inside the region of homology (i.e. inside the targeting cassette) and a second fluorescence marker gene differing from the first marker gene outside the region of homology. Moreover, it was found that with a gene targeting vector containing a positive selection marker (e.g. neomycin) inside the targeting cassette and a fluorescent marker outside the region of homology (e.g. ZsGreen) identification of both (stably transfected) non-targeted and correctly targeted ES cells via optical detection was possible.

Furthermore, stably gene targeting vector transfected ES cells have been automatically analyzed by fluorescence (Cytocon™300; Evotec Technologies) and plated as single cells (Cytocon™ Single Cell Fraction Collector; Evotec Technologies). Both high resolution analysis and single cell plating can be combined in Image Activated Cell Sorting (Elektra, Evotec Technologies). Amplification of clonally

plated ES cells enriched for gene targeting events allowed for the first time the isolation of gene targeting events without need for any manual picking or sub-cloning of cells. Furthermore, the contact free procedure and gentle cell handling kept targeted ES cells germline-competent and allowed the production of chimeric and, by tetraploid complementation (Nagy, A. et al., Proc. Natl. Acad. Sci. USA 90:8424-8428 (1993); Eggen, K. et al., Proc. Natl. Acad. Sci. USA 98(11):6209-14 (2001)) fully ES cell derived animals.

Altogether, this method allows for the first time automated gene-targeting from transfection of the gene targeting vector to establishment of totipotent targeted ES clones.

The invention thus provides

(1) a method for isolation of primary mammalian after homologous recombination comprising

(a) transfecting starting primary cells with a homologous targeting vector comprising

(A) a homologous targeting cassette which comprises (i) a functional DNA segment and a positive selection marker or a functional DNA segment and a first detectable marker, and (ii) two DNA segments homologous to the integration site within the genome of the primary mammalian cells flanking (i); and

(B) an expression cassette harboring a DNA sequence coding for a second detectable marker different from said first detectable marker, said expression cassette (B) being connected with said homologous targeting cassette (A) so as to allow distinction between targeted and non-targeted cells;

(b) manual or automatic identification and/or isolation of cells containing the positive selection marker or the first detectable marker; and

(c) manual or automatic identification and/or isolation of cells as homologous recombinants by the absence of the second detectable marker;

(2) a method for preparing transgenic tissues, organs and/or multi-cell organisms which comprises utilizing the method as defined in (1) above; and

(3) a vector for targeted homologous recombination of eukaryotic cells as defined in (1) above.

The vector utilized in the homologous recombination process of (1) above allows direct or indirect distinction between correctly targeted and non-targeted cells.

Indirect distinction can be achieved via addition of non-toxic compounds to the cells or via metabolic or enzymatic action of the gene product (detectable marker) encoded by the expression cassette.

By utilizing one or more fluorescence markers as detectable markers, the method (1) of the invention

- a) allows to optically distinguish random DNA integrants (light emission) and HR eukaryotic cells, including ES cells (no light emission) (Fig. 2);
- b) enhances isolation and reduces variability of recovery of HR ES cells in individual experiments since the addition of compounds to the culture medium is avoided;
- c) allows selective, automated distinction of growing transfected ES cell colonies by plating of ES cells, transformed with such light emission targeting vectors, in a way such that preferentially single cell clones grow within a spatially separated region (1 well of a multiwell-dish) and subsequent fluorescence analysis (such ES clones can then automatically be expanded and taken to molecular analysis (Fig. 2));
- d) allows selective, automated sorting of growing transfected ES cells for the absence of the first detectable marker and/or the presence of the second detectable marker and/or by Image Activated Cell Sorting (i.e. fluorescence).

The method of the invention has the following advantages over current technology:

- (i) The detection strategy allows selective identification of gene targeted ES clones without exposure of ES cells to potentially toxic chemicals, protein or mutagens usually used for negative selection;
- (ii) automated optical selection of clones enriched for homologous recombination event by applying gentle sorting machinery equipment such as but not limited to Cytocon™300 & Cytocon™ Single Cell Fraction Collector (Evotec Technologies) or Elektra (Evotec Technologies);
- (iii) delivery of isolated single cells in 1 well of a multiwell dish which retain the capability of clonal growth;

- (iv) identification of a higher percentage of clones showing homologous recombination compared to standard procedures (i.e. positive-negative selection, positive-positive selection) due to specific Promoter and selectable marker configurations (WO 03/002725);
- (v) automated expansion, isolation and molecular analysis of gene targeted ES clones with proven characteristics of totipotency as demonstrated by the generation of chimeric and tetraploid mice;
- (vi) significant reduction of both manual labour and associated costs for gene targeting experiments.

Short Description of the Figures

Fig. 1: Selection mechanism in homologous recombination: A: targeting vector; B: homologous genomic DNA; C: homologous recombined DNA, presence of first marker, absence of second marker.

Fig. 2: Scheme of manual and automated gene targeting process employing novel detectable markers (solid arrows indicate experiments described in this application).

Fig. 3: Generation of targeted ES cells. Scheme of the gene targeting strategy. Both constructs, (i) the targeting vector for positive and negative selection (pOMP1) (3A), and (ii) the targeting vector for positive and optical detection (pOMP3) (3B) were used for insertion of the neo gene into the *Rosa26* locus by homologous recombination as depicted. E: EcoRV; X: XbaI.

(3C) Southern blot analysis of genomic DNA from ES cells transfected with constructs described in (Fig 4A) and (Fig 4B). The DNA of ES cells was digested with EcoRV and hybridized with probe 1 specific for the *rosa26* locus. The sizes of the wt and the targeted allele are 11.5 kb and 2.5 kb, respectively. The picture shows the analysis of identified non-fluorescent clones.

(3D) Southern blot analysis of genomic DNA from ES cells transfected with pOMP3 and automatically sorted and plated using Cytocon™ 300.

Fig. 4: Growth of targeted fluorescent and non-fluorescent (marked with arrow) ES clones.

Fig. 5: Enrichment of gene targeting frequency in non-fluorescent ES clones.

Fig. 6 shows schematically the functional segments of the *rosa26* locus.

Fig. 7 shows the vector pOMP3.

Fig. 8 shows the vector POMP1.

Fig. 9 shows schematically the modified *rosa26* locus.

Fig. 10 shows schematically the CAGGS promoter followed by an intron.

Detailed Description of the Invention

"Targeted vector" or "vector for targeted homologous recombination" according to the present invention is a DNA construct that contains sequences "homologous" to endogenous chromosomal nucleic acid sequences which flank a desired genetic modification. The flanking homologous sequences direct the targeting vector to the specific chromosomal location within a genome by virtue of the homology that exists between the flanking homologous sequences and the corresponding endogenous sequences and introduce the desired genetic modification (i.e. a functional DNA segment) by a process referred to as "homologous recombination".

"Homologous" means two or more nucleic acid sequences that are either identical or similar enough that they are able to hybridize to each other or undergo intermolecular exchange.

A "flanking DNA" is a segment of DNA that is collinear with and adjacent to a particular point of reference.

"Gene targeting" is the modification of an endogenous chromosomal locus by the insertion into, deletion of or replacement of the endogenous sequence via homologous recombination using a targeting vector.

The "homologous targeting cassette" is that part of the targeted vector that carries the desired genetic modification, the flanking homologous sequences and optionally functional sequences such as selectable or detectable markers.

The "expression cassette" is that part of the targeted vector that enables distinction of non-targeted and correctly targeted stably transfected cells by determining its presence or absence in the transfected cellular system. The two cassettes are "connected" through covalent linkage of an arbitrary number of nucleotides.

A "detectable marker" only allows for the detection and subsequent isolation of a transfected cell, whereas a "selectable marker" or "selection marker" provides for a selection advantage of the cell containing said marker over cells devoid of said marker and therewith allows for the enrichment of rare transfected cell expressing the marker from the majority of treated cells in the population.

"Multi-cell organisms" according to the present invention refers to non-human mammals. The "transgenic tissues" or "transgenic organs" are derived from the multi-cell organisms referred to above.

A "non-human organism" or "non-human mammal" is an organism or mammal that is not normally accepted by the public as being human.

As set forth above, the vector used for targeting the primary mammalian cells in the method of (1) above has the following properties: The homologous targeting cassette (A) comprises (i) a functional DNA segment and a positive selection marker or a functional DNA segment and a first detectable marker, and (ii) two DNA segments homologous to the integration site within the genome of the primary mammalian cells flanking said functional DNA segment and said positive selection marker or said functional DNA segment and said first detectable marker. The expression cassette (B) comprises a DNA sequence coding for a second detectable marker different from said first detectable marker.

It is preferred that the expression cassette (B) comprises a gene coding for said first detectable marker under the control of a promoter (e.g. constitutive, inducible, etc.) active in primary mammalian cells. Preferably the promoter is a constitutive promoter including, but not limited to, CAGGS (Niwa, H.K. et al., Gene 108(2):193-9 (1991)), CMV (Bi, J.X. et al., Biotechnol. Bioeng. 81(7):848-54 (2003)), c-fos (Bi, J.X. et al., Biotechnol. Bioeng. 81(7):848-54 (2003)), PGK (Adra, C.N. et al., Gene 60:65-74 (1987)), SV40 (Southern, P.J., Berg, P., J. Mol. Appl. Genet. 1:327-341 (1982)), elongation factor 1alpha (eF1alpha), RNA polymerase II (Soriano, P. et al., J. Virol. 65(5):2314-9 (1991)) and TK.

Moreover, it is preferred that said first and/or second detectable marker are a non-toxic, directly or indirectly detectable compound or the marker(s) are gene(s) expressing membrane bound protein(s) or protein(s) with membrane anchoring signal sequence, requiring that such proteins are not expressed in the transfected cell of interest. Such membrane bound proteins can be identified by binding to fluorescence or otherwise labeled detectable antibodies.

Preferably the detectable marker is a compound which is detectable by colorimetric (Lowry method, Bradford reaction (Comassie blue), sulforhodamine B, β -galactosidase (lacZ), placental alkaline phosphatase (PAP)), fluorescence, bioluminescence such as luciferases from phengodid beetles, which can produce green to red bioluminescence Viviani, V.R., Cell Mol. Life Sci. 59(11):1833-50 (2002)) phosphorescence detection methods or the like (for review of probes see (Haughland, R.P., Molecular Probes. Handbook of fluorescent probes and research chemicals, 172-180, 221-229 (1992-1994); Freshney, R.I., Culture of animal cells, 3rd ed., Wiley & Sons, Inc. (1994)). Most preferred is the use of a fluorescence protein/peptide.

In another preferred embodiment, the first and/or second detectable marker, preferably the second detectable marker, is a membrane bonded protein selected from CD1a, CD2, CD3, CD4, CD5, CD6, CD8, CD11, CD14, CD15, CD16, CD19, CD20, CD22, CD25, CD27, CD30, CD33, CD34, CD43, CD45, CD56, CD61, CD62, CD69, CD71, CD90, CD105, CD117, CD123, CD133, CD138, BDCA-2, BDCA-3, BDCA-4, CRTH2, ErbB-2, Ep-CAM (Human Epithelial Antigen HEA), etc. Particularly preferred is that the positive selection marker is neomycine and the second selectable marker is selected from ZsGreen CD4 and CD8, or the first and second selectable markers are ZsGreen and HcRed.

According to the invention the expression cassette may be positioned 5' or 3' of a homologous targeting cassette. If the homologous targeting cassette harbors at least two regions of target gene homology which differ in length, the position of the expression cassette is preferably connected with the longer region of homology.

According to the invention the primary mammalian cells include pluripotent and totipotent cells and preferably are embryonic stem cells (hereinafter shortly referred

to as "ES cells") (excluding human ES cells if this is not patentable). Preferably the primary mammalian cells are derived from a rodent, such as mouse, rat, etc. Particularly preferred are mouse ES cells.

In a more preferred embodiment the vector of (1) above the functional DNA segment of the targeting cassette is a DNA sequence encoding a gene of interest, or a part of a gene of interest, or is a functional DNA sequence which can be converted into such DNA sequence encoding a gene of interest, or intronic sequence, or is a regulatory functional DNA sequence, including, but not limited to, splice receptor or splice donor or acceptor sequences, and recombinase recognition site(s) (RRS(s)).

The functional DNA segment may further comprise gene expression control elements including, but not limited to, ubiquitous or tissue specific promoter, either constitutive or inducible, a polyadenylation signal, intron sequences, recombinase recognition site(s), enhancer recognition site(s), and matrix attachment region(s) (MAR).

The positive selection marker according to the present invention may be a DNA sequence encoding a protein conferring resistance against cell poison, including, but not limited to, neomycine (Beck, E. et al., *Gene* 19(3):327-36 (1982)), hygromycine (Bernard, H. U. et al., *Exp. Cell Res.* 158(1):237-43 (1985); Gordano, T. J., McAllister, W. T., *Gene* 88(2):285-8 (1990); Santerre, R. F. et al., *Gene* 30(1-3):147-56 (1984)), puromycine N acetyltransferase (*puro*) (de la Luna, S. et al., *Gene* 62(1):121-6 (1988); de la Luna, S., Ortin, J., *Methods Enzymol.* 216:376-85 (1992)), *E. coli* xanthine guanosine phosphoribosyltransferase (*gpt*) (Spring, K. J. et al., *Biochim. Biophys. Acta* 1218(2):158-62 (1994)), blasticidin (Kobayashi, K. et al., *Agric. Biol. Chem.* 55(12):3155-7 (1991); Kimura, M. et al., *Mol. Gen. Genet.* 242(2):121-9 (1994)), hypoxanthine-guanosine phosphorybosyltransferase (*Hprt*) (Albertini, R. J. et al., *Nature* 316(6026):369-71 (1985); Lester, S. C. et al., *Somatic Cell. Genet.* 6(2):241-59 (1980)), herpes simplex virus type 1 thymidine kinase (*hsvTK*) (Borrelli, E. et al., *Proc. Natl. Acad. Sci. USA* 85(20):7572-6 (1988)), and adenine phosphoribosyltransferase (*aprt*) (Lester, S. C. et al., *Somatic Cell. Genet.* 6(2):241-59 (1980)) - HPRT, TK and *aprt* in positive selection schemes require ES cells with mutations in their endogeneous *Hprt* or *TK* loci - or may be a

DNA sequence, conferring superior metabolic properties to the cells, as compared to non-transfected (starting) cells, including the utilization of xanthine, adenine, etc.

According to the present invention the first and/or second marker are DNA sequences encoding a protein allowing direct optical detection, such as by detection of fluorescence, chemiluminescence or the like, or indirect optical detection, such as a colorimetric assay or the like, provided that said second detectable marker differs from said first detectable marker in order to allow separate detection. It is particularly preferred that the first and/or second detectable marker peptide is a fluorescence protein including, but not limited to, green fluorescence protein (GFP) and homologous, analogous or color variants thereof, and Reef Coral Fluorescent Proteins (RCFPs) such as AmCyan, ZsGreen, ZsYellow, DsRed, AsRed and HcRed (Clontech, Palo Alto, CA). Particularly preferred are the fluorescent peptides mentioned in the following Table 1 (Labas, Y.A. et al., PNAS 99(7):4256-4261 (2002)):

Table 1: Summary of family of GFP-like proteins

Protein ID	Original ID	GenBank accession #	Genus species
amajGFP	amFP486	AF 168421	<i>Anemonia majano</i>
dsfrGFP	DsFP483	AF 168420	<i>Discosoma striata</i>
clavGFP	CFP484	AF 168424	<i>Clavularia sp.</i>
GFP		M 62653	<i>Aequora victoria</i>
cgigGFP		AY 037776	<i>Condylactis gigantea</i>
hcriGFP		AF 420592	<i>Heteractis crispa</i>
ptilGFP		AY 015995	<i>Ptilosarcus sp.</i>
rmueGFP		AY 015996	<i>Renilla muelleri</i>
zoanGFP	zFP560	AF 168422	<i>Zoanthus sp.</i>
asulGFP	asFP499	AF 322221	<i>Anemonia sulcata</i>
dis3GFP		AF 420593	<i>Discosoma sp.3</i>
dendGFP		AF 420591	<i>Dendronephthya sp.</i>
mcavGFP		AY 037769	<i>Montastraea cavernosa</i>
rfloGFP		AY 037772	<i>Ricordea florida</i>
scubGFP1		AY 037767	<i>Scolymia cubensis</i>
scubGFP2		AY 037771	<i>Scolymia cubensis</i>
zoanYFP		AF 168423	<i>Zoanthus sp.</i>
DsRed	drFP583	AF 168419	<i>Discosoma sp.1</i>
dis2RFP	dsFP593	AF 272711	<i>Discosoma sp.2</i>
zoan2RFP		AY 059642	<i>Zoanthus sp.2</i>
mcavRFP		AY 037770	<i>Montastraea cavernosa</i>
rfloRFP		AY 037773	<i>Ricordea florida</i>

asulCP	asCP	AF 246709	<i>Anemonia sulcata</i>
hcrlCP	hcCP	AF 363776	<i>Heteracis crista</i>
cgigCP	cpCP	AF 363775	<i>Condylactis gigantea</i>
cpasCP	cpCP	AF 383155	<i>Condylactis parsiflora</i>
gtenCP	gtCP	AF 383156	<i>Goniopora tenuidens</i>

It is moreover preferred that the positive selection marker or the second detectable marker are flanked by RRSs in order to allow removal of the marker by means of a recombinase after the proper recombination product was established.

The flanking homologous DNA segments may have a length of 0.1 to 20 kb, preferably 0.5 to 10 kb. The actual flanking DNA sequences depend on the locus of integration, suitable sequences are those homologous to the Rosa26, HPRT, beta-actin, GAPDH locus or the like of eukaryotic cells.

In a particularly preferred vector of the invention the first detectable marker is ZsGreen and the constitutive promoter is the CAGGS promoter. The targeting cassette in such a vector may contain a positive selection marker as defined above or a second fluorescence marker differing from the first marker. Suitable marker combinations out of the ones mentioned in Table I are known in the art. Reef Coral Fluorescent Proteins (RCFPs) such as AmCyan, ZsGreen, ZsYellow, DsRed, AsRed and HcRed (Clontech, Palo Alto, CA) each emit a distinct wavelength. The corresponding cDNAs were isolated from nonbioluminescent reef corals (class Anthozoa), and have been optimized for bright emission, fast chromophore maturation, and codon optimized for increased expression in mammalian cells. Because of their distinct spectra RCFPs can be used in combination to visualize multiple events simultaneously.

In a most preferred vector the homologous sequences of the targeting cassette are homologous to the mouse Rosa26 locus and the selection marker is a loxP flanked neomycin resistance gene (see bp 1287 to 3233 of SEQ ID NO:6), preferably said vector has the sequence of SEQ ID. NO. 6.

The method (1) of the invention comprises, but is not limited to, the following steps:

1. Generation of a gene targeting construct harbouring one or more expression cassettes coding for non-identical optical markers, e.g. two fluorescence molecules (ZsGreen, AmCyan1 or ZsYellow1 or DsRed2, DsRedExpress or AsRed2, or HcRed1; Clontech, Palo Alto, CA) under control of a constitutive CAGGS promoter (Niwa, H. et al., Gene, 108(2):193-9 (1991)).
2. Gene targeting by homologous recombination in embryonic stem (ES) cells with this vector.
3. Identification of preferentially homologous recombined ES clones harbouring the second detectable marker, not harbouring the first detectable marker.
4. Isolation of such clones by
 - (a) manual picking of non-fluorescent clones under fluorescence light emission source;
 - (b) limited dilution plating of electroporated ES cells in multi-well cell culture dishes at a pre-determined concentration to enable preferential growth of single clones/well;
 - (i) manual sub-culture and molecular analysis of identified non-fluorescent colonies by current protocols;
 - (ii) automated sub-culture and molecular analysis of identified non-fluorescent colonies;
 - (c) automated cell sorting, preferably by image activated cell sorting of preferentially homologous recombined ES cells expressing or not expressing the detectable fluorescent marker and plating of sorted cells at single cell density in multi-well tissue culture plates.

Said automated sorting is preferably performed with one of the following devices A to C with the particular settings indicated:

(A) Cytocon™ 300 (Evotec Technologies) equipped with Cytocon™ Single Cell Fraction Collector (Evotec Technologies). Parameters are preferably as follows:

- I) Software: Cytocon™ software "SWITCH"
- II) Cytocon™ 300 sorter chip mounted on an Olympus IX-50 fluorescence microscope equipped with a 100 W lamp and fluorescein filter set (BP510-550, DM570, BA590). Single cells are preferably caged within the field cage using 5 V at 800 kHz.
- III) Cytocon™ Fluidics system connected to the Cytocon™ 300 sorter chip,

consisting of

- a) precision syringe pump
 - rate pump1: 1:1000-1500 pl/s
 - rate pump 2: -488 pl/s
- b) Fluidic block
- c) Injector
 - a. Hamilton syringe, 5 – 10 µl injection volume

I) Cytocon™ chip driver, settings as follows:

- a. Mode Zero: 605
- b. Mode One: 4.88

Detached fluorescent & nonfluorescent ES cells are preferably resuspended in an isotonic buffered solution (Cytocon-bufferII™) at a density of 1×10^5 cells/500µl.

Preferably for reduction of electrolytes buffering is achieved with but not limited to carbohydrates such as glucose or inositol. Exposure to such buffer is preferably minimized to less than 30 minutes, more preferably less than 20 minutes.

Fluorescence is assessed by eye (Olympus IX-50 fluorescence microscope equipped with a 100 W lamp and fluorescein filter set (BP510-550, DM570, BA590); objective 10x/0.30 Ph1 UplanF1).

Single cells showing no fluorescence are flushed out of the Sorter Chip with ES cell culture medium and deposited with aid of Cytocon™ Single Cell Fraction Collector (Evotec Technologies) into separate wells of a 96 well microtiter plate pre-filled with ES cell culture medium with the following parameters: Sheath flow: 800000 pl/s (0.8µl/s)

(B) Elektra (Evotec Technologies)

Sorter chip mounted on an Olympus IX-50 fluorescence microscope equipped with a 100 W lamp and fluorescein filter set (BP510-550, DM570, BA590). Single cells are preferably caged within the field cage using 6V, 700 kHz. Fluidics system connected to the sorter chip, consisting of precision syringe pump (rates pump1: 1:1000-1500 pl/s, pump 2: -488 pl/s). Injector with Hamilton syringe (10 µl injection volume)

Detached fluorescent & nonfluorescent ES cells are preferably resuspended in an isotonic buffered solution (Cytocon-bufferII™) at a density of 1×10^5 cells/500µl.

Preferably for reduction of electrolytes buffering is achieved with but not limited to carbohydrates such as glucose or inositol. Exposure to such buffer is preferably minimized to less than 30 minutes, more preferably less than 20 minutes.

Fluorescence is assessed by eye (Olympus IX-50 fluorescence microscope equipped with a 100 W lamp and fluorescein filter set (BP510-550, DM570, BA590); objective 10x/0.30 Ph1 UplanF1).

Single cells showing no fluorescence are flushed out of the Sorter Chip with ES cell culture medium and deposited into separate wells of a 96 well microtiter plate pre-filled with ES cell culture medium with the following parameters: Sheath flow: 800000 pl/s (0.8 µl/s)

(C) Magnetic Bead Separation is preferably performed with the BD™ IMag Cell Separation System (BD Bioscience), more preferably with the MACS® Technology (Milteny Biotec) based on the use of MACS MicroBeads, MACS Columns and MACS Separators; or separation is performed with the ClinIMACS® Cell Selection System.

The method (2) of the invention for preparing transgenic tissues, organs and/or multi-cell organisms (except for human being or human tissues and organs, if such are not patentable under the respective regulations) is exemplified by the following steps:

(a) Generation of genetically modified organisms by

(aa) ES cells. Such cells can be used to create genetically modified rats or mice by standard blastocyst injection technology (Robertson, E. J. Editor, IRL Press: Oxford, UK, 71-112 (1987)), aggregation (Wood, S. A. et al., Nature 365(64411):87-89 (1993)), tetraploid blastocyst injection (Wang, Z. Q. et al., Mech. Dev. 62(2):137-45 (1997)), nuclear transfer and cloning (Wakayama, T. et al., Proc. Natl. Acad. Sci. USA 96(26):14984-9 (1999)), ES cells derived from other organism such as rabbits (Wang, Z. Q. et al., Mech. Dev. 62(2):137-45 (1997); Schoonjans, L. et al., Mol. Reprod. Dev. 45(4):439-43 (1996)), chickens (Pain, B. et al., Development 122(8):2339-48 (1996)), or other species;

(bb) modified protoplasts to generate genetically modified plants;

(cc) nuclear transfer from modified eukaryotic cells to oocytes to generate cloned organisms with modified allele (Wakayama, T. et al., Proc. Natl. Acad.

Sci. USA 96(26):14984-9 (1999); Baguisi, A. et al., Nat. Biotechnol. 17(5):456-61 (1999); Wilmut, I. L. et al., Nature 385(6619):810-3 (1997); Wilmut, I. L. et al., Reprod. Fertil. Dev. 10(7-8):639-43 (1998); Wakayama, T., Protein, Nucleic Acid Enzyme 45(13 Suppl.):2005-14 (2000); Wakayama, T. et al., Nature 394(6691):369-74 (1998); Rideaut, W. M., Nat. Genet. 24(2):109-10(2000); Campbell, K. H. et al., Nature, 380(6569):64-6 (1996));

(dd) cell-fusion to transfer the modified allele to another cell, including transfer of engineered chromosome(s), and uses of such cell(s) to generate organisms carrying the modified allele or engineered chromosome(s) (Kuroiwa, Y. et al., Nat. Biotechnol. 18(10):1086-90 (2000)).

(b) Cloning: The techniques used to construct DNA vectors described herein are standard molecular biology techniques well known to the skilled molecular biologist (Sambrook, J. Fritsch, E. F. a.M., Molecular Cloning: a Laboratory Manual, 2nd ed., vols. 1, 2 and 3, Cold Spring Harbor, NY, USA, Wiley Interscience, NY (1989)).

(c) Gene modifications including

(aa) deletion of coding sequences, gene segments, or regulatory elements;

(bb) alterations(s) of coding sequence, gene segments, or regulatory elements including substitutions, additions, and fusions (e.g. epitope tags or creation of bifunctional proteins);

(cc) insertion of new coding regions, gene segments, or regulatory elements, such as those for selectable marker genes or reporter genes or putting new genes under endogenous transcriptional control;

(dd) creation of conditional alleles, e.g. by introduction of loxP sites flanking the region to be excised by Cre recombinase (Abremski, K., Hoess, R., J. Biol. Chem. 259(3):1509-14 (1984)) or FRT sites flanking the region to be excised by FLP recombinase (Andrews, B. J. et al., Cell 40(4):795-803 (1985); Meyer-Leon, L. et al., Cold Spring Harb. Symp. Quant. Biol. 49:797-804 (1984)); or

(ee) replacement of coding sequences or gene segments from one species with orthologous coding sequences from a different species, e.g. replacing a murine genetic locus with the orthologous human genetic locus to engineer a mouse where that particular locus has been 'humanized'.

(d) Introduction of vectors into eukaryotic cells: Using standard methodology, such as transfection mediated by calcium phosphate, lipids, or electroporation (Sambrook, J. Fritsch, E. F. a.M., Molecular Cloning: a Laboratory Manual, 2nd ed., vols. 1, 2 and 3, Cold Spring Harbor, NY, USA, Wiley Interscience, NY (1989)).

(e) Selection: By exposure to selection agents, depending on the selectable marker gene that has been engineered into the vector. If the selectable marker is the neomycin phosphotransferase (neo) gene (Beck, E. et al., Gene 19(3):327-36 (1982)) then cells that have taken up the vector can be selected in G418-containing media; cells that do not have the vector will die whereas cells that have taken up the vector will survive (Freshney, R. I., Culture of animal cells, 3rd ed., Wiley & Sons, Inc. (1994)). Other suitable selectable markers include any drug that has activity in eukaryotic cells, such as hygromycin B Santerre, R. F. et al., Gene 30(1-3):147-56 (1984); Bernard, H. U. et al., Exp. Cell. Res. 158(1):237-43; Giordano, T. J., McAllister, W. T., Gene 88(2):285-8 (1990)), Blasticidin (Izumi, M. et al., Exp. Cell. Res. 197(2):229-33 (1991)) and other which are familiar to those skilled in the art.

The invention is further described by reference to the following examples, which is, however, not to be construed so as to limit the invention.

Example

As example for optical identification of ES clones, the fluorescent molecule ZsGreen gene (Clontech) was chosen. The Rosa locus of the mouse (Fig.6, SEQ ID NO:3) was chosen for homologous recombination. The targeting strategy is outlined in Figures 3A, B and C and D.

1. Rosa Targeting Vector: A 129 Sv/Ev-BAC library (Incyte Genomics) was screened with a probe against *exon2* of the *Rosa26* locus (amplified from mouse genomic DNA using Rscreen1s (GACAGGACAGTGCTTGTTTAAGG; SEQ ID NO:1) and Rscreen1as (TGA CTACACAATATTGCTCGCAC; SEQ ID NO:2)). PCR conditions were as follows: 95°C, 2 min, followed by 30 cycles: 95°C; 30 s; 60°C, 30 s; 72°C, 30 s; 72°C, 7 min; followed by 20°C, 2 min. Out of the identified BAC clone a 11 kb EcoRV subfragment was inserted into the HindII site of pBS. Two fragments, a 1 kb SacII/XbaI fragment (SEQ ID NO:4) and a 4 kb XbaI-fragment (SEQ ID NO:5) were

used as homology arms and inserted into a vector consisting of a FRT-flanked neomycin resistance gene and a PGK-TK-pA expression cassette for negative selection (SEQ ID NO:13).

2. pOMP1 and pOMP3 vector Construction: For construction of pOMP1 (Fig. 8 and SEQ ID NO:7), the FRT-flanked neomycin resistance gene was replaced by a loxP-flanked neomycin resistance gene (SEQ ID NO:9). The ZsGreen gene was placed under the control of the CAGGS promoter (SEQ ID NO:11) followed by a synthetic intron (pMultilink CAGGS-zsGreen; Fig. 10, SEQ ID NO:10). This expression cassette was used for a replacement of PGK-TK-pA of pOMP1 resulting in pOMP3 (Fig. 7 and SEQ ID NO:6). pOMP1 and pOMP3 were used for the targeting experiments (Figures 3A, 3B). For ES cell electroporation, vectors were linearized at the I-SceI restriction site.

3. Introduction of pOMP vectors into ES cells

The ES cell line Art4.12 (Eggan, K. et al., Nat. Biotechnol., 20(5):p. 455-9 (2002)) was grown on mitotically inactivated feeder layer (Mitomycin C (Sigma M-0503)) comprised of mouse embryonic fibroblasts in medium composed of 1x DMEM high Glucose (Invitrogen 41965-062), 4 mM Glutamin (Invitrogen 25030-024), 1x Non Essential Amino Acids (Invitrogen 11140-035) 1mM Sodium Pyruvate (Invitrogen 11360-039), 20 mM Hepes (Invitrogen 15630-056), 0.1 mM β -Mercaptoethanol (Invitrogen 31350-010), 2×10^6 μ /l Leukemia Inhibitory Factor (Chemicon ESG 1107) and 15% fetal bovine serum (PAN 1302-P220821).

Linearized vectors pOMP1 and pOMP 3 were introduced into the cells by electroporation using a Gene Pulser with Capacitance Extender (Biorad).

Rapidly growing cells were used on the first day following the last passage. Upon trypsinization (Invitrogen 25200-056) cells were resuspended in PBS (Invitrogen 20012-019) and preplated for 25 min on gelatinized 10 cm plates to remove unwanted feeder cells. The supernatant was harvested, ES cells were washed once in PBS and counted (Neubauer hemocytometer). 10^7 cells were mixed with 30 μ g of I-SceI linearized vector in 800 μ l of transfection buffer (20 mM Hepes, 137 mM NaCl, 15 mM KCl, 0.7 mM Na_2HPO_4 , 6 mM Glucose 0.1 mM β -Mercaptoethanol in H_2O) and electroporated using a Biorad Gene Pulser with Capacitance Extender set on 250 V and 500 μ F. Electroporated cells were seeded at a density 0.25×10^7 cells per 10 cm

tissue culture dish onto a previously prepared layer of neomycin-resistant inactivated mouse embryonic fibroblasts (MEF). For automated selection of non-fluorescent clones (experiment 6), cells were seeded on 6 cm gelatine-coated tissue culture dishes without MEF layer.

4. Manual identification of ES cells containing a targeted disruption of the Rosa locus

48 h after electroporation, the medium was replaced on all dishes by medium containing 250 µg/ml Geneticin (Invitrogen 10131-019) for positive selection of G418 resistant ES clones.

5 day after electroporation, ½ of the dishes containing ES cells electroporated with vector pOMP1 were negatively selected by addition of 2 µM Gancyclovir (Cymeven®, Roche)

On day 8 after electroporation ES colonies were isolated as follows:

Medium was replaced by PBS and the culture dishes were placed on the stage of a binocular (Nikon SMZ-2B). Using low magnification (25 x) individual ES clones of undifferentiated appearance were removed from the surface of the culture dish by suction into the tip of a 20 µl pipette (Eppendorf). The harvested clones were placed in individual wells of 96 well plates containing 30 µl of 2.5% Trypsin (Invitrogen). After disruption of clones by pipetting with a multichannel pipette (Eppendorf), cells were seeded onto feeder containing 96 well plates with pre-equilibrated complete ES-medium.

For pOMP1 positive selected clones (Geneticin) and positive/negative selected clones (Geneticin and Gancyclovir) were harvested as described above.

For pOMP3 only positive selected clones were isolated either at random as described above, or the ES clones were placed under a binocular (Leica MZFL III) with fluorescent light source (GFP3 Filters set) and bright-field illumination. At dim light, clearly identifiable non-fluorescent ES clones (Figure 4) were isolated as described above.

Cells were grown for 3 days with daily medium changes and then splitted 1 : 2 on gelatinized (Sigma G-1890) 96 well plates. 3 days after splitting cells were lysed, genomic DNA was prepared and analysed by Southern blot.

5. Identification of non-fluorescent transfected cells by plate-reader analysis

Electroporated cells were seeded at a density of 1 cell per well in a 96 multi-well plate (Nunc, Order #136101), coated with a previously prepared layer of neomycin-resistant inactivated mouse embryonic fibroblasts.

48 h after electroporation, the medium was replaced on all dishes by medium containing 250 µg/ml Geneticin (Invitrogen 10131-019) for positive selection of G418 resistant ES clones.

5 days upon transfection individual wells were scored by plate-reader analysis

- a) by O.D. measurement for the presence of single ES clones
- b) by fluorescence measurement (GFP3 filter set; Leica, Bensheim, Germany) for the location of non-fluorescent clones.

Identified non-fluorescent ES clones were expanded and taken to molecular analysis by Southern blotting (see Fig. 3A, 3B) according to standard procedures.

6. Image Activated Cell Sorting for the Analysis of Transfected Cells and Single Cell Plating

48 h after electroporation, the medium was replaced on all dishes by medium containing 250 µg/ml Geneticin (Invitrogen 10131-019) for positive selection of G418 resistant ES clones.

On day 7 or 8 after electroporation ES colonies were automatically sorted by Cytocon™ 300 (Evotec Technologies) based on the presence or absence of fluorescent light emission of single clones. Non-fluorescent ES cells were automatically plated employing a Cytocon™ Single Cell Fraction Collector (Evotec Technologies) in single entities of a 96-well plate.

Transfected ES cells were enzymatically removed from by trypsin addition (Invitrogen 25200-056). After centrifugation, cells were resuspended in Cytocon buffer II at a density of 1×10^5 cells/500µl. Five to ten µl of 1:5 in buffer II diluted cells were applied to Cytocon™ 300 using a Hamilton syringe. Sorting and single cell plating were undertaken with the following Cytocon™ 300 settings: Generator modes 0 605; 1 4.88; pump rates pump1: 1:1000-1500 pl/s; pump 2: -488 pl/s; "sheat flow": 800000 pl/s (0.8µl/s). A Cytocon™ Sorter Chip was mounted on a fluorescent microscope and single cells were caged within the field cage. Fluorescence was assessed by eze and non-fluorescent cells were automatically

seeded with the aid of the Cytocon™ Single Cell Fraction Collector (Evotec Technologies) into single entities (individual well of 96-well tissue culture plate (Greiner Cellstar), coated with MEF cells and pre-equilibrated complete ES-medium).

Clones were grown for 7 days with one medium change and then replated and cultured 1:1 onto a second 96-well tissue culture plate. Clones were subsequently expanded, aliquots cryoconserved and genomic DNA prepared and molecularly analysed by Southern blot.

7. Injection of ES cells into diploid and tetraploid blastocysts and generation of chimeric and ES mice

Embryo culture was carried out in microdrops on standard bacterial petri dishes (Falcon) under mineral oil (Sigma). Modified CZB media (Chatot et al, Supra (?)) was used for embryo culture unless otherwise noted. Hepes buffered CZB was used for room temperature operations.

Chimeras were generated as described (Hogan, B., Beddington, R., Costantini, F. & Lacy, E. eds. Manipulating the Mouse Embryo, a Laboratory Manual. 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor, Laboratory Press, 1994). Briefly, Balb/C host embryos were harvested at dpc 3.5 from the uterus of superovulated Balb/c OlaHsd females (Harlan Netherlands) mated with Balb/c OlaHsd males (Harlan Netherlands). For microinjection, 5-6 blastocysts were placed in a drop of DMEM with 15% FCS under mineral oil. A flat tip, piezo actuated microinjection-pipette with an internal diameter of 12-15 µm was used to inject 20 - 30 ES cells into each blastocyst. After recovery, 8 injected blastocysts were transferred to each uterine horn of 2.5 days post coitum, pseudopregnant NMRI females that had been mated with vasectomized males. Litters were controlled, and pups alive by that time were counted as surviving pups. ES cell contribution was judged by coat color chimerism.

Production of mice by tetraploid embryo complementation has been previously described (Eggan et al., PNAS 98:6209-6214 (2001)). After administration of hormones, superovulated B6D2F1 females were mated with B6D2F1 males. Fertilized zygotes were isolated from the oviduct and any remaining cumulus cells

were removed with hyaluronidase. After overnight culture, two-cell embryos were electrofused to produce one cell tetraploid embryos using a CF150-B cell fusion instrument from BLS (Budapest, Hungary) according to the manufacturers instructions. Embryos that had not undergone membrane fusion within 1 hour were discarded. Embryos were then cultured in vitro to the blastocyst stage. For microinjection, 5-6 blastocysts were placed in a drop of DMEM with 15% FCS under mineral oil. A flat tip, piezo actuated microinjection-pipette with an internal diameter of 12-15 μm was used to inject 15 ES cells into each blastocyst. After recovery, ten injected blastocysts were transferred to each uterine horn of 2.5 days post coitum, pseudopregnant NMRI females that had been mated with vasectomized males. Litters were controlled, and pups alive by that time were counted as surviving pups.

8. Results

(a) Effect of fluorescence on ES cell growth

Expression of ZsGreen (pOMP3) did not negatively influence ES cell colony formation. The mean number of resistant ES colonies per 1×10^7 ES cells transfected with the ZsGreen expression vector pOMP3 is similar to the number of ES colonies transfected with pOMP1 (see Table 2).

Table 2: Fluorescence does not negatively effect clonal ES cell growth

	2 nd marker gene	2 nd Selection	# clones per 1×10^7 transfected cells
pOMP1	None	none	3500
POMP1	Thymidine Kinase	Gancyclovir	950
pOMP3	ZsGreen	none	3366

(b) Selective isolation of non-fluorescent ES colonies.

15,5% of all colonies in the pOMP3 gene targeting experiment were non-fluorescent. Isolation of non-fluorescent ES colonies under fluorescent light-source as described above, led to significant enrichment of such ES colonies. By this method 99,3% (all but 1 clone) were non-fluorescent when monitored upon isolation see Table (3).

Table 3: Enrichment of isolation of non-fluorescent ES colonies in pOMP3 experiments.

		Isolated (%)
Random isolation	Fluorescent	84,5
Random isolation	Non-fluorescent	15,5
Selective isolation	Fluorescent	0,7
Selective isolation	Non-fluorescent	99,3

(c) HR frequency in pOMP targeting experiments

The analysis of non-fluorescent ES cells for HR events in the case of pOMP3 targeting resulted in a targeting frequency of 6,49% \pm 0,5% in 3 independent experiments (Table 4).

This is a 8fold higher frequency if compared to targeting of pOMP1 without counterselection. Compared to pOMP1 gene targeting plus counterselection with Gancyclovir (3,02% HR), optical isolation was 2,2fold more efficient.

Table 4: Enrichment of gene targeting frequency in non-fluorescent ES clones. (see also Fig. 5)

Construct	Selection	HR - clones/analyzed	HR frequency (%)
pOMP1	None	6/708	0,85
pOMP1	TK/GanC	10/331	3,02
POMP3	fluorescent	0/452	0,00
POMP3	Non-fluorescent	16/239	6,69

(d) HR frequency in Automated pOMP targeting experiments

Transfected, sorted and individually seeded cells attached and divided at a frequency of 34%. Transfected, sorted and individually seeded cells attached and divided at a frequency of 34%. 9 individual clones were analyzed and HR clones identified at a frequency of 4%.

Table 5: Frequency of gene targeting in FACS sorted and single-cell plated ES clones.

Construct	Single cell plating efficiency	HR clones/analyzed	HR frequency
POMP3	34%	1/25	4%

(e) The automated gene targeting procedure does not affect competence of mouse production of wild-type "Art4.12" ES cells and HR ES clone "OMP 3 Cyto 1-B6". Cytocon™ 300 sorted ES cells produce highly coat colour chimeric mice (see Table 6), and, via tetraploid complementation, solely ES clone derived mice (see Table 7).

Table 6: Production of Chimeric mice by diploid Injection.

	Art 4.12 Parental	Targeted ES clone OMP 3 Cyto 1-B6
# <i>Injected diploid</i>	61	52
# Pups born	14	13
# Chimeric pups	11	13

Table 7: Production of ES mice by tetraploid Injection.

	Art 4.12 Parental	Targeted ES clone OMP 3 Cyto 1-B6
# <i>Injected tetraploid</i>	153	60
# ES mice born	1	3

Claims

1. A method for isolation of primary mammalian after homologous recombination comprising

(a) transfecting starting primary cells with a homologous targeting vector comprising

(A) a homologous targeting cassette which comprises (i) a functional DNA segment and a positive selection marker or a functional DNA segment and a first detectable marker, and (ii) two DNA segments homologous to the integration site within the genome of the primary mammalian cells flanking (i); and

(B) an expression cassette harboring a DNA sequence coding for a second detectable marker different from said first detectable marker, said expression cassette (B) being connected with said homologous targeting cassette (A) so as to allow distinction between targeted and non-targeted cells;

(b) manual or automatic identification and/or isolation of cells containing the positive selection marker or the first detectable marker; and

(c) manual or automatic identification and/or isolation of cells as homologous recombinants by the absence of the second detectable marker.

2. The method of claim 1, wherein

(i) the expression cassette (B) comprises a gene coding for said second detectable marker under the control of a promoter active in primary mammalian cells, preferably the promoter is a constitutive promoter; and/or

(ii) said first and/or second detectable marker is a non-toxic, directly or indirectly detectable compound, preferably is a compound being detectable by colorimetric, fluorescence, chemiluminescence, phosphorescence detection methods, is a membrane bound or protein with membrane anchoring signal sequence or the like, most preferably is a fluorescent peptide; and/or

(iii) the expression cassette (B) is positioned 5' or 3' relative to the homologous targeting cassette; and/or

(iv) the primary mammalian cells are pluripotent or totipotent cells, preferably are ES cells; and/or

(v) the primary mammalian cells are derived from a rodent such as mouse, rat, etc. preferably the primary mammalian cells are mouse ES cells.

3. The method of claim 1 or 2, wherein

(i) the functional DNA segment of the targeting cassette (A) is a DNA sequence encoding a gene of interest, or is a functional DNA sequence which can be converted into such DNA sequence encoding a gene of interest, or intronic sequence, or is a regulatory functional DNA sequence, including, but not limited to, splice receptor of splice donor or acceptor sequences, and recombinase recognition site(s); and/or

(ii) the functional DNA segment may further comprise gene expression control elements including, but not limited to, ubiquitous or tissue specific promoter, either constitutive or inducible, a polyadenylation signal, intron sequences, recombinase recognition site(s) (RRS), enhancer recognition site(s), and matrix attachment region(s) (MAR); and/or

(iii) the positive selection marker is a DNA sequence encoding a protein conferring resistance against cell poison, including, but not limited to, neomycine, hygromycine, puromycine, histidinol and bleomycine, is a DNA sequence, conferring superior metabolic properties to the cells, including the utilization of xanthine, adenine etc., or the like; and/or

(iv) the first and/or second detectable marker are DNA sequences encoding a protein allowing direct optical detection, such as by detection of fluorescence, chemiluminescence or the like, or indirect optical detection, such as a colorimetric assay or the like, provided that said second detectable marker differs from said first detectable marker in order to allow separate detection; and/or

(v) the second detectable marker is a membrane bound or protein with membrane anchoring signal sequence; and/or

(vi) the positive selection marker and the first detectable marker are flanked by one or more RRSs; and/or

(vii) the flanking homologous DNA segments have a length of 0.1 to 20 kb, preferably 0.5 to 10 kb; and/or

(viii) the flanking DNA sequences are homologous to the Rosa26, HPRT, beta-actin, GAPDH locus of the eukaryotic cells.

4. The method of claim 2 or 3, wherein

(i) the first and/or second detectable marker peptide is a fluorescence protein including, but not limited to, green fluorescence protein (GFP) and modifications thereof, the constitutive promoter of the expression cassette is selected from GAGGS, CMV, PGK, TK, preferably the fluorescence protein is ZsGreen and the constitutive promoter is the CAGGS promoter; and/or

(ii) the first and/or second detectable marker is a protein including, but not limited to, CD1a, CD2, CD3, CD4, CD5, CD6, CD8, CD11, CD14, CD15, CD16, CD19, CD20, CD22, CD25, CD27, CD30, CD33, CD34, CD43, CD45, CD56, CD61, CD62, CD69, CD71, CD90, CD105, CD117, CD123, CD133, CD138, BDCA-2, BDCA-3, BDCA-4, CRTH2, ErbB-2, and Ep-CAM (Human Epithelial Antigen HEA) and preferably the positive selection marker is neomycine and the second selectable marker is selected from ZsGreen CD4 and CD8, or the first and second selectable markers are ZsGreen and HcRed.

5. The vector of claim 4, wherein the homologous sequences are homologous to the mouse Rosa26 locus, and preferably said vector has the sequence of SEQ ID. NO.

6.

6. The method according to any one of claims 1 to 5, wherein the identification or isolation is performed automatically, preferably

(i) the isolation comprising effected by fluorescence activated cell sorting (FACS); and/or

(ii) the isolation buffer is an isotonic buffered solution, preferably a reduction of electrolytes buffering is achieved with but not limited to carbohydrates such as glucose or insitol and/or

(iii) the cells to be sorted remain in the buffer for a time shorter than 30 minutes, preferably shorter than 20 minutes.

7. The method according to any one of claims 1 to 6 which further

(i) plating of transfected cells into standard tissue culture vessels; and/or

(ii) limited dilution plating of transfected cells into multiwell plates, including 96well, 384 well and 1536 well plates, in order to allow preferential growth and analysis of single clones in individual wells/plate, and optionally

(iii) subsequent detection and isolation of transfected cells plated as described

above by manually or automatically selecting for the absence of the first detectable marker and/or the presence of the second detectable marker.

8. The method of claim 7 wherein the presence of the first and/or absence of the second marker is analyzed by (FACS) mechanism, whereby the isolated sorted cells are re-plated as described in claim 7.

9. A method for preparing transgenic tissues, organs and/or multi-cell organisms which comprises utilizing the method as defined in claims 1 to 8.

10. The method of claim 9, wherein the transgenic multi-cell organism is a non-human mammal, most preferably mouse, and said method comprises modifying an ES cell as defined in claims 1 to 9.

11. The method of claim 10 which further comprises one or more of the steps
(i) injecting isolated ES cells in or aggregating isolated ES cells with diploid and/or multiploid preimplantation embryos, preferably blastocysts, or
(ii) injecting in/fusion of nuclei from ES cells with enucleated non-fertilized eggs (nuclear transfer); and/or
(iii) generating transgenic non-human embryos and/or animals.

12. A vector for targeted homologous recombination of eukaryotic cells as defined in claims 1 to 5.

-1/13-

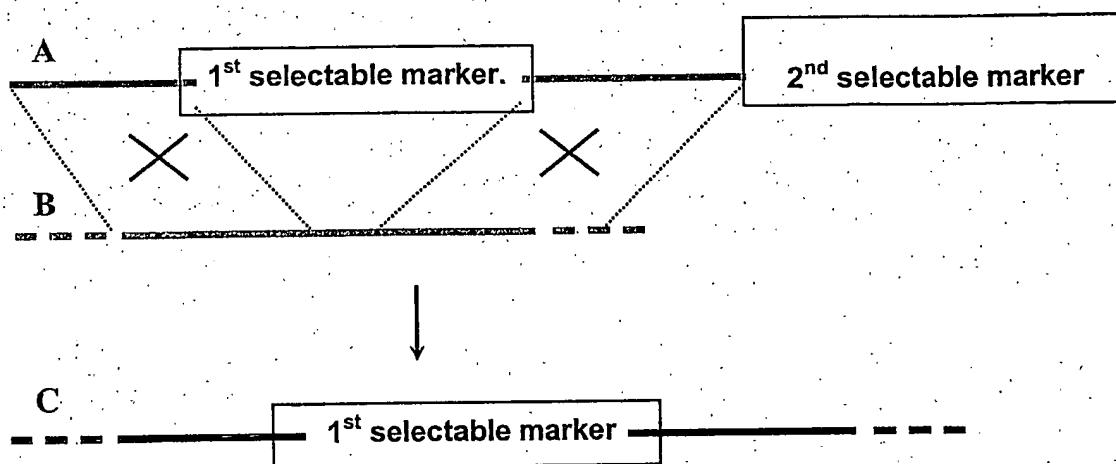


Fig. 1

-2/13-

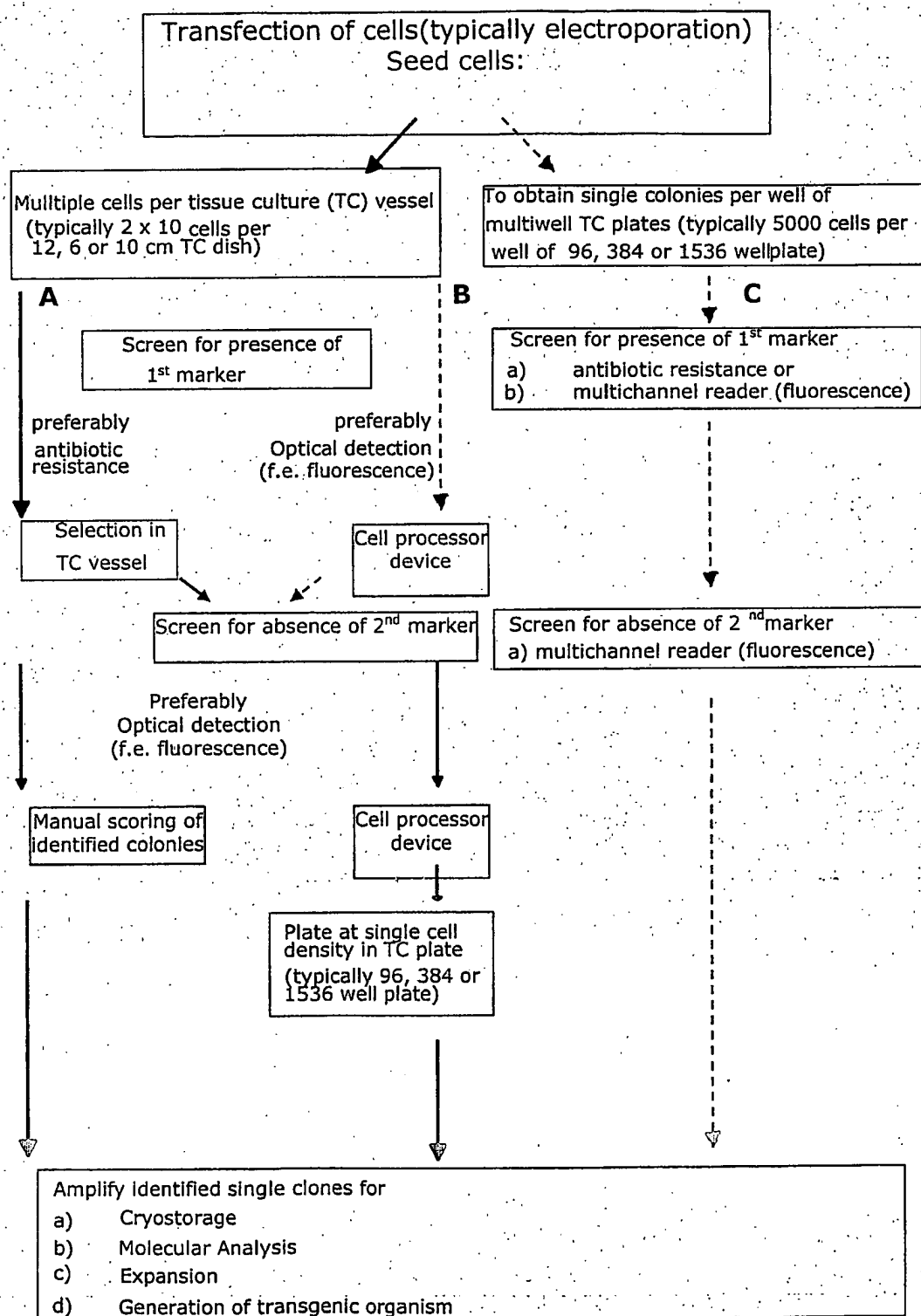


Fig.2

-3/13-

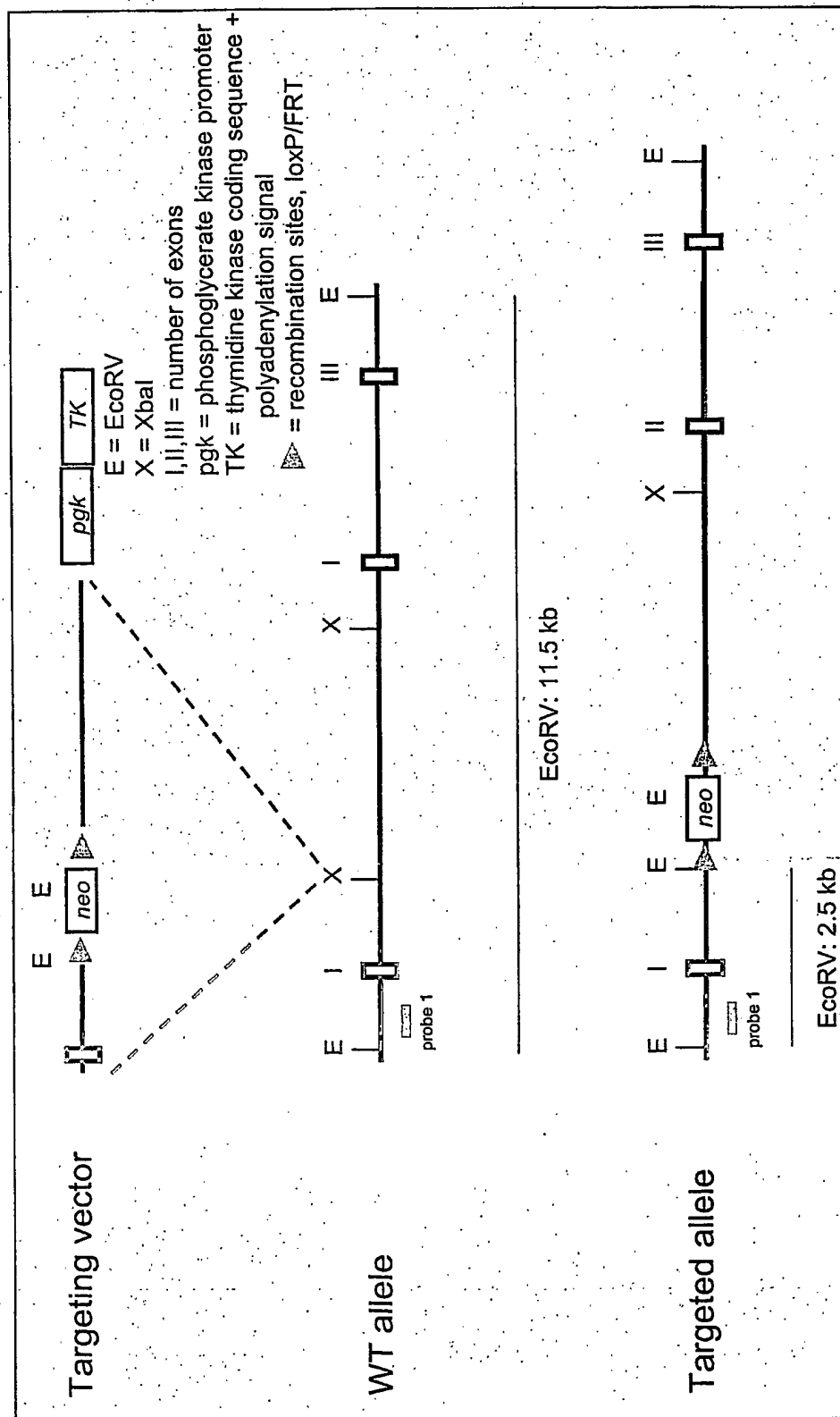


Fig. 3a

-4/13-

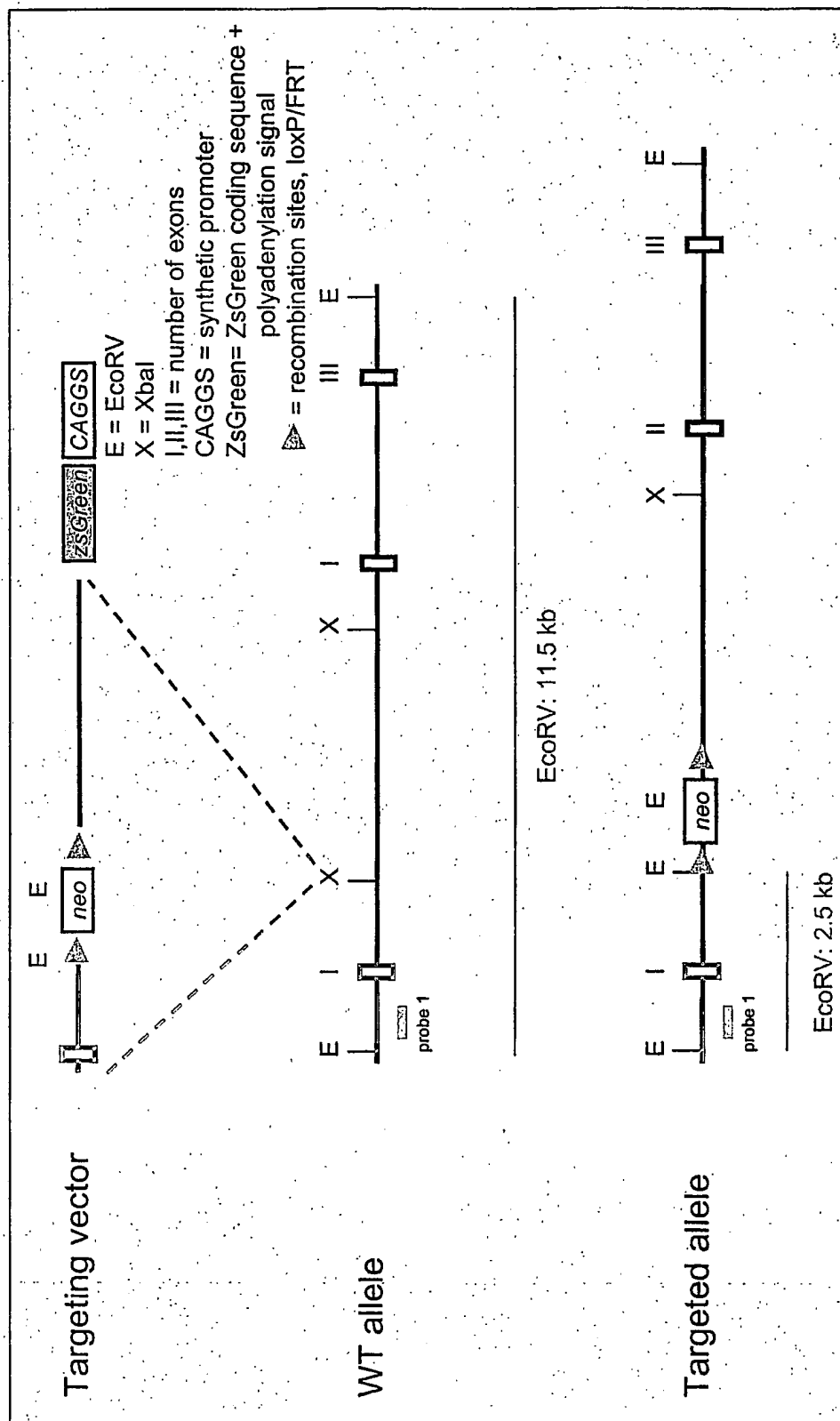


Fig. 3b

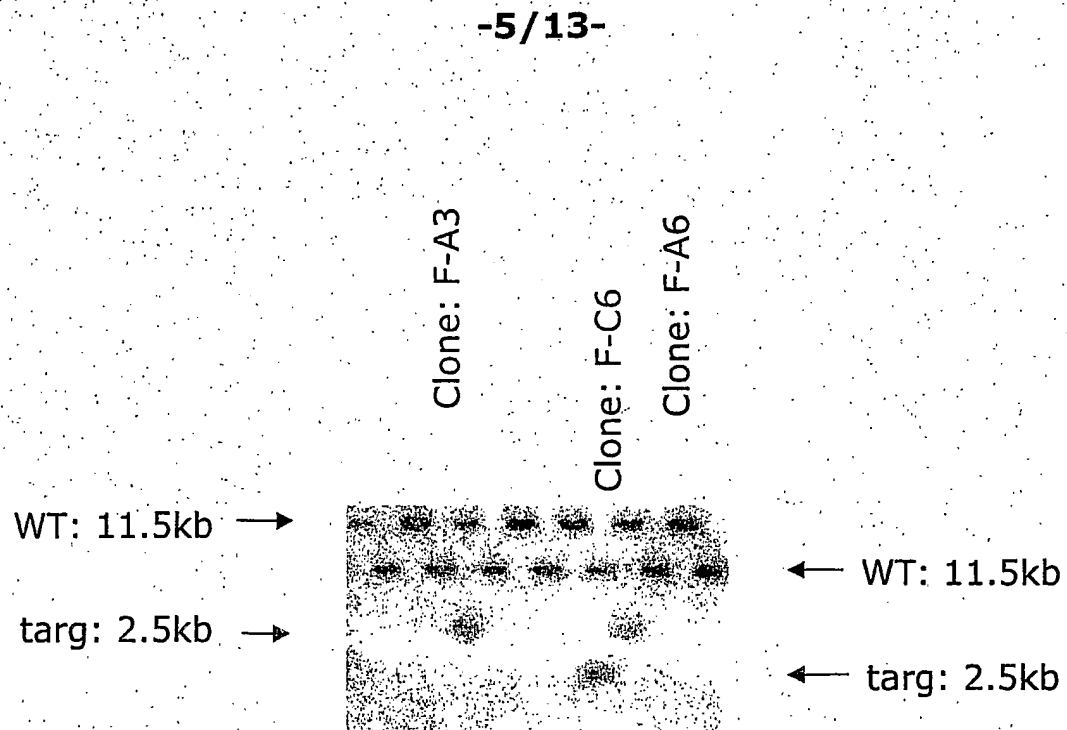
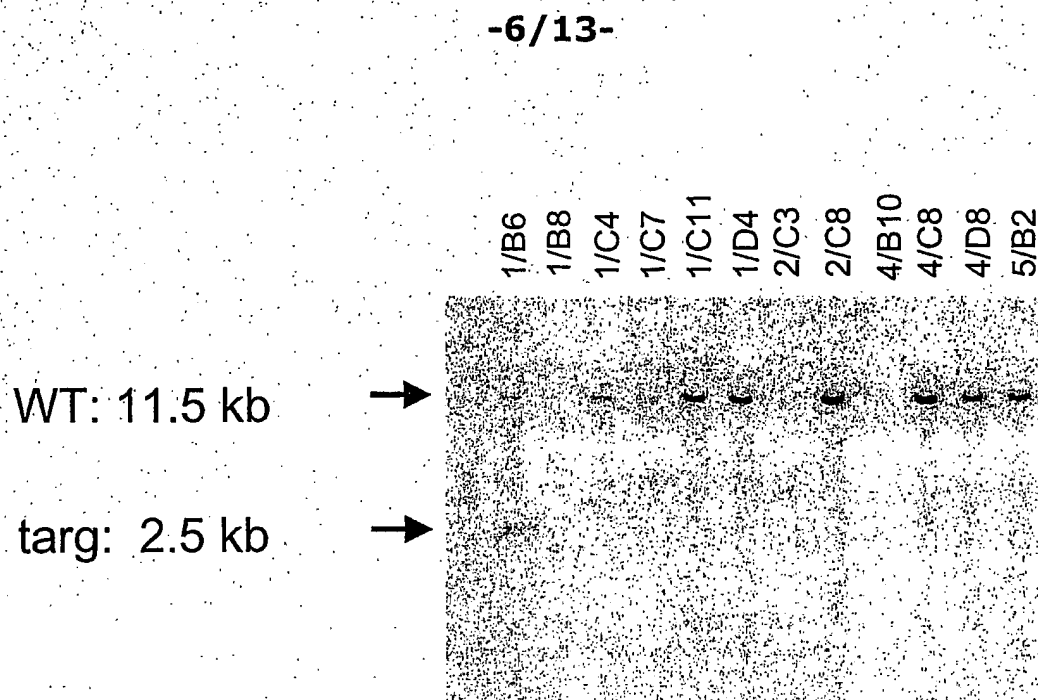


Fig.3c

**Fig.3d**

- 7/13 -

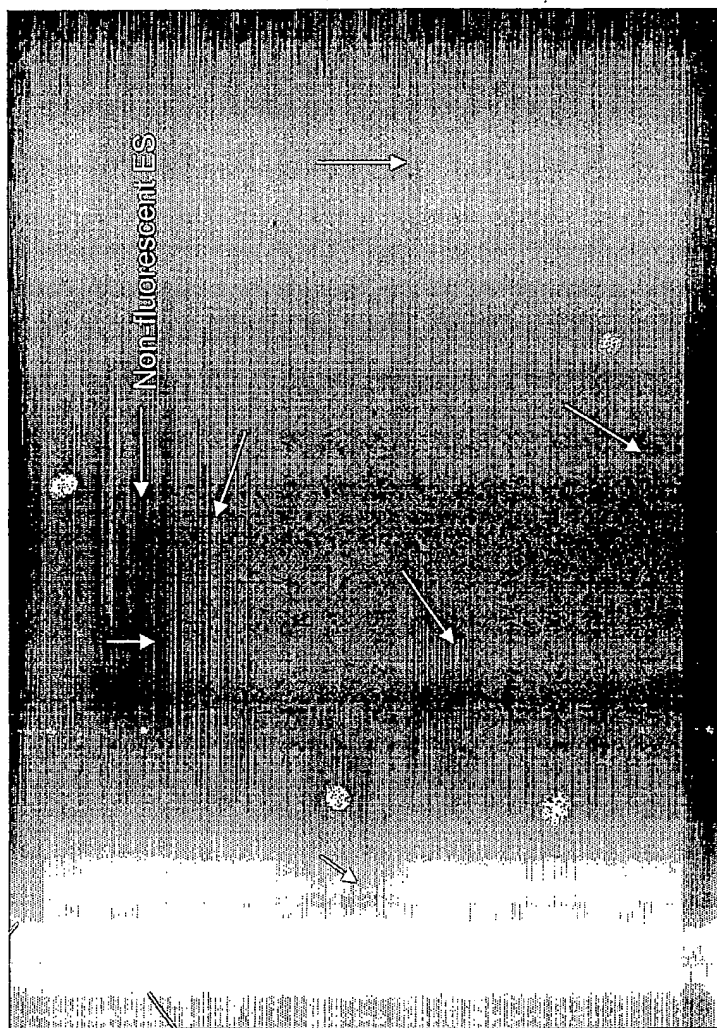
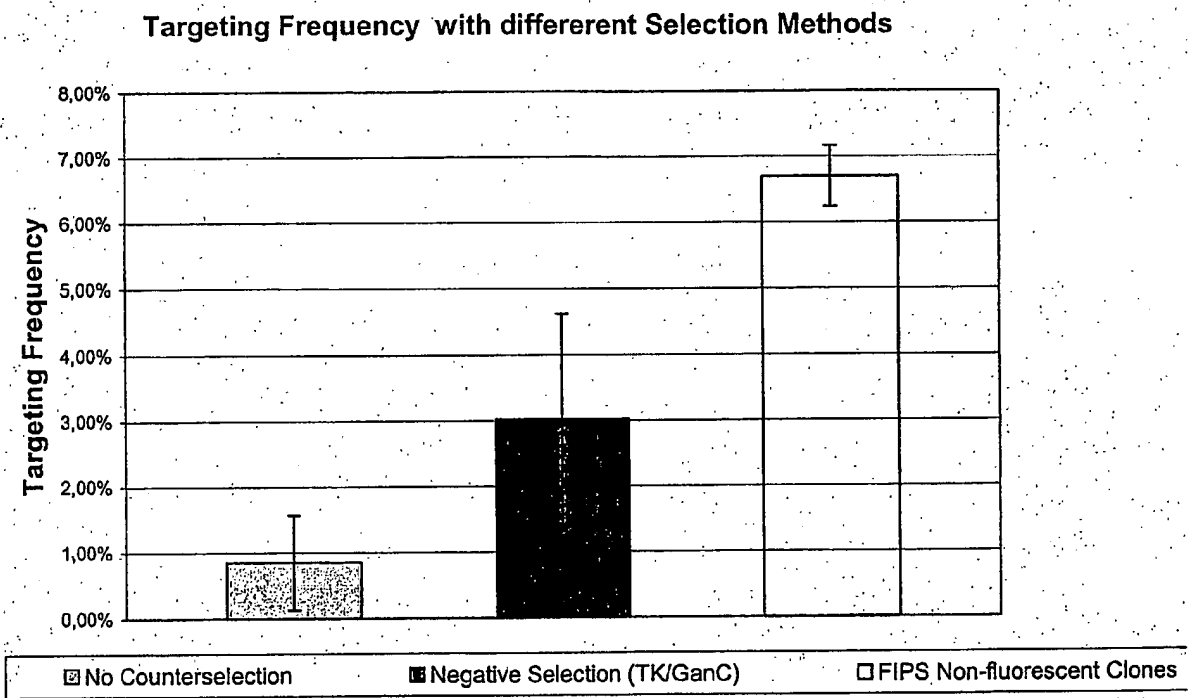


Fig. 4

-8/13-**Fig.5**

- 9/13-

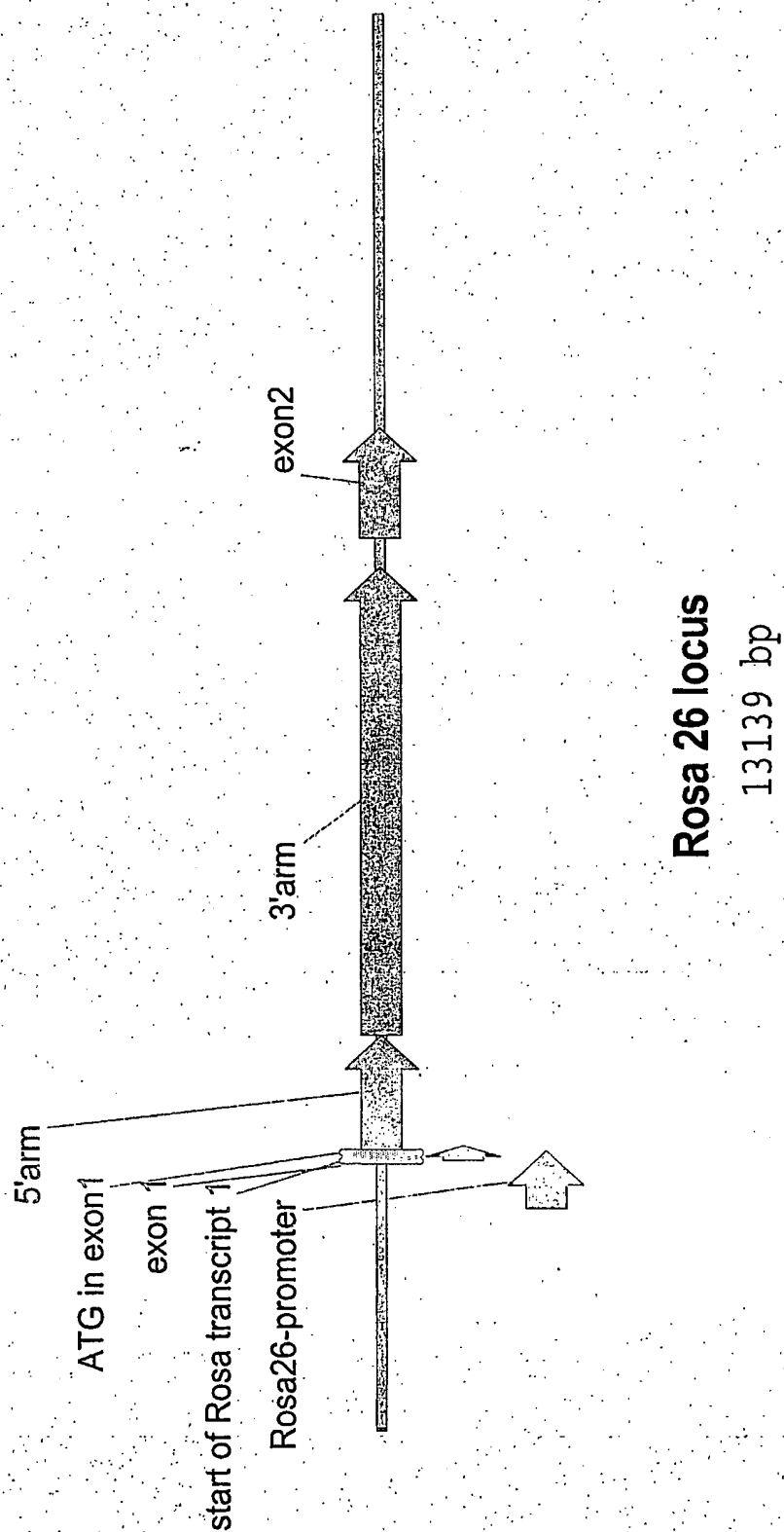


Fig. 6

- 10/13 -

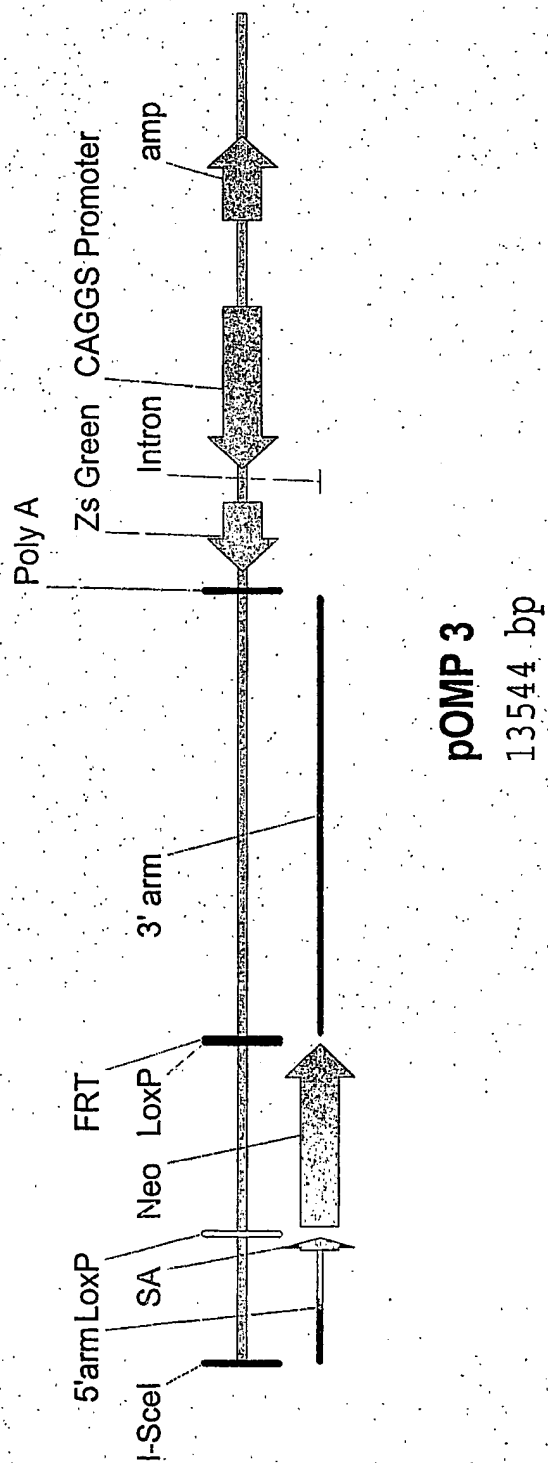


Fig. 7

- 11/13 -

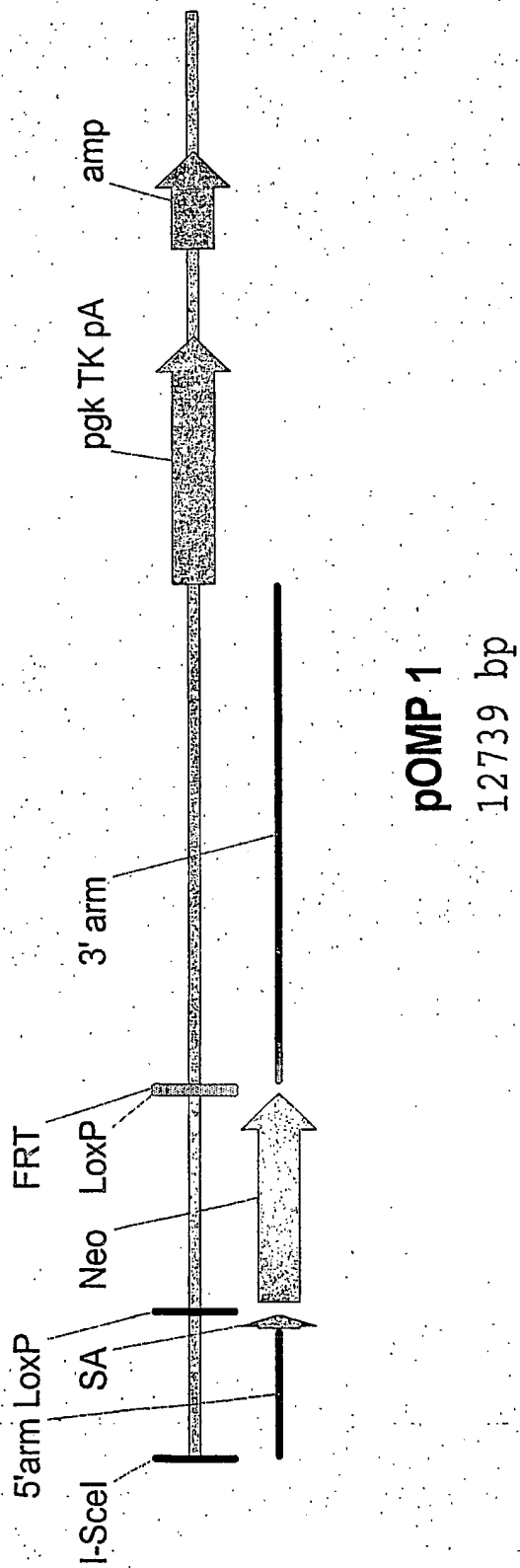


Fig. 8

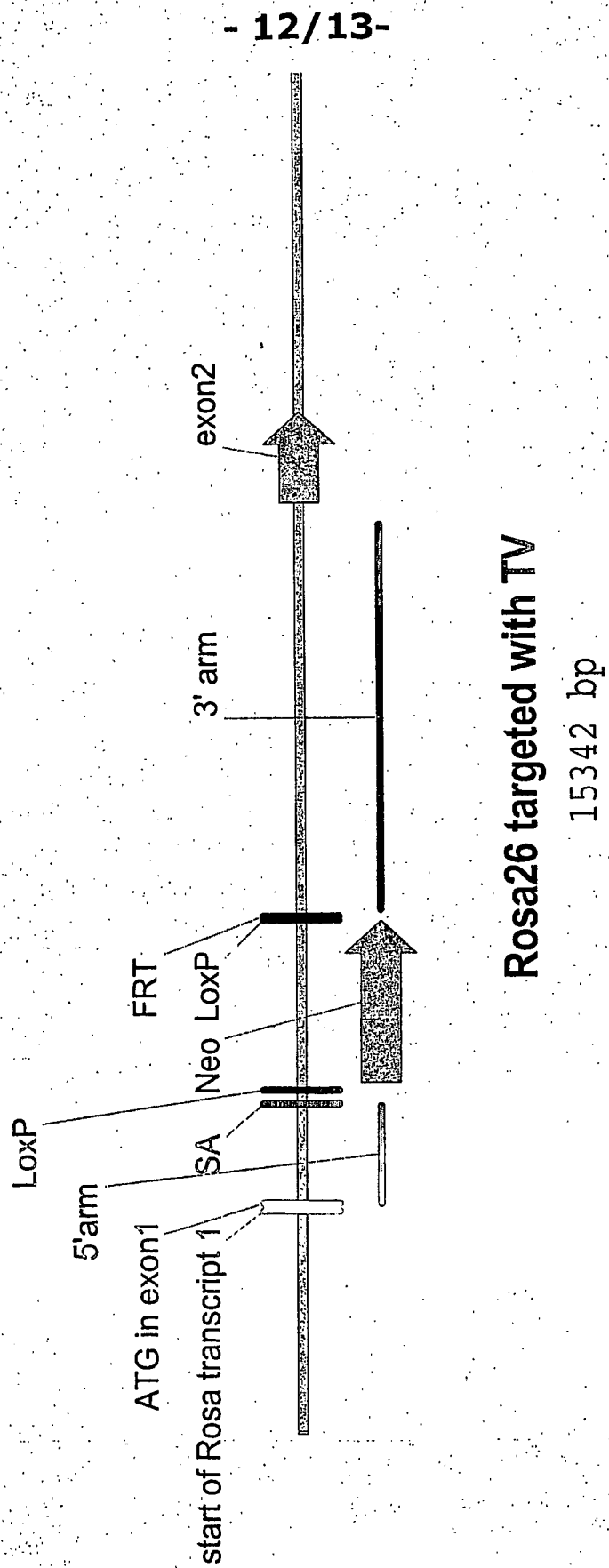


Fig. 9

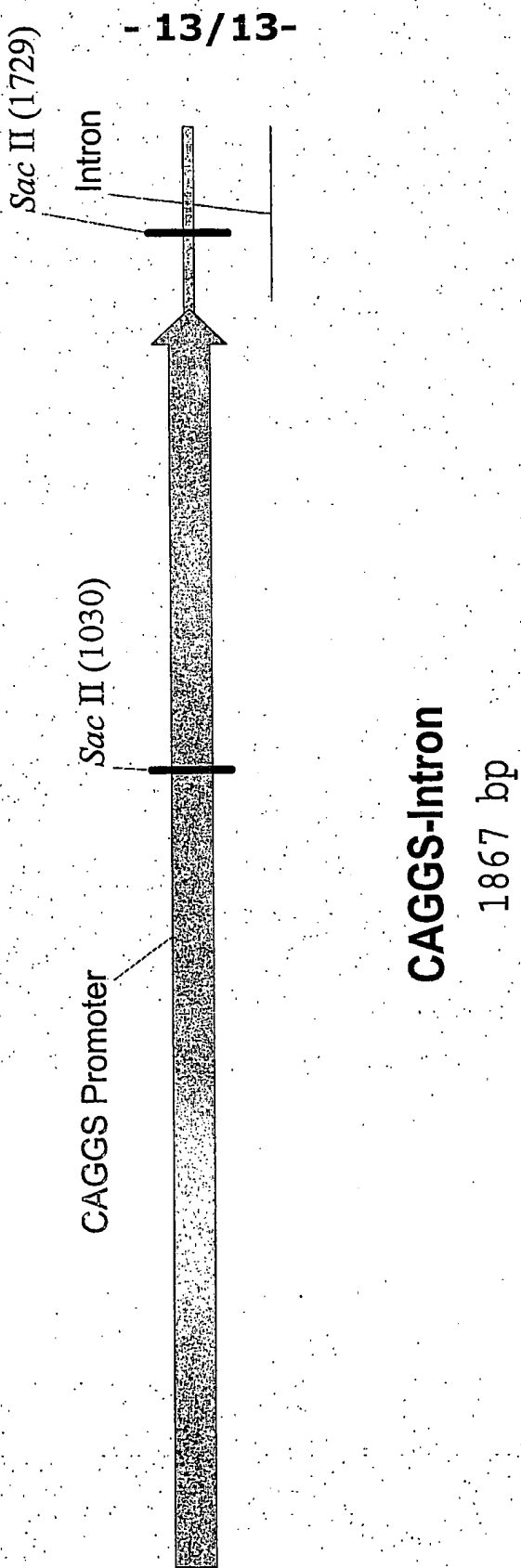


Fig. 10

SEQUENCE LISTING

<110> Artemis Pharmaceuticals GmbH

<120> Automated Gene-Targeting Using Non-Toxic Detectable Markers

<130> 040104wo JH/BM

<140>

<141>

<150> 03002424.4

<151> 2003-03-05

<150> US 60/446,921

<151> 2003-03-12

<160> 13

<170> PatentIn Ver. 2.1

<210> 1

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer
Rsceenls

<400> 1

gacaggacag tgcttggtta agg

23

<210> 2

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer
Rscreenlas

<400> 2

tgactacaca atattgctcg cac

23

<210> 3

<211> 13139

<212> DNA

<213> Mus musculus

<220>

<223> Description: Rosa 26 locus sequence

<400> 3

aagcttctca cgtagcaacc agagctccag agccagcagc tgctgcccgc ttgtatactc 60
actcctgtga tccaacacag gagcaacott ttctttaccc cacccccact tcttaacaca 120
cttttttttg gggggggggg gggaacaagt gtcctatgct ggaaggattg gaactatgct 180
tttagaaaag aacaatccta aggtcacttt taaattgagg tctttgattt gaaaatcaac 240
aaataccaaa ttccaaatat tcgttttaat taaaccagca atgtggatat aagcattaag 300
ttttagtttt aaaaagggtca attttccaaa cattcagcaa tcatatttaa atttacagct 360
aggaacaaga gccttgggtc atgtcctacc aaagaacata actcaatatt ctacacatga 420
caatctgaat aaccttaaa cctctaattc cataacaggc cacaaatttt ggacagagaa 480
ctaattgatcc tcttgagaaa actggaagaa atccagggaa aagaaattcc tgtgtcctcc 540
aaactcagaa atctctaatt atgtcagtat tctctgcttt agtcctaggt cagattgcac 600

acatctaaaa	taacctctta	aagttttcct	cctagcgacc	taaaccatta	ttaatatcaa	660
attaaccatc	aaaacacttt	cctctcaata	tgctgcacac	aaacctctc	ctggaacctc	720
ctccatctgg	atcctcccca	atcaaaagta	taggtattta	acataataagc	aaggaagtaa	780
tgtaaacatg	accttggtca	caaatatgtc	atctaaaaac	aatttagtca	aggtatggag	840
gaaattcgag	aacctgaatc	tttttaagta	ttttgagcac	aggaacaatt	ggcaaaaagg	900
atccagggtat	agacaaaacc	cagagcccag	agctctgggc	gaaaaatgag	ttgctggtga	960
agacgttaca	caagtaacat	gagaaagcag	aaaatgcagg	tcattccacgc	acctctgacc	1020
caggccagca	gggcgggctg	cagcatcagt	acacaggaga	aagatcctta	ttcctaagaa	1080
tgagaaaggc	aaaggcgccc	gatagaataa	attagcatag	aaggggcttt	cccaggagtt	1140
aaaactttcc	ttctgagcga	ttacctacta	aaaccagggc	ttttgccac	taccatttac	1200
ctaggatctt	ggcttgacag	gattcatagg	ggcatatccc	tccccctctt	ctttagagtc	1260
gttcttaaaa	gatcgctctc	cacgcccctg	gcagggaata	cgacaaaatc	tggtcaatt	1320
ccaggctaga	accotacaaa	ttcagcaggg	atattcgcaag	gatactgggg	catagccac	1380
agggagtcca	agaatgtgag	gtgggggtgg	cgaaggtaat	gtctttggtg	tgggaaaagc	1440
agcagccatc	tgagatagga	actggaaaac	cagaggagag	gcgttcagga	agattatgga	1500
ggggaggact	gggccccccac	gagcgaccag	agttgtcaca	aggccgcaag	aacaggggag	1560
gtggggggct	cagggacaga	aaaaaaagta	tggtattttt	gagagcaggg	ttgggaggcc	1620
tctcctgaaa	aggtataaaa	cgtggagtag	gcaataccca	ggcaaaaagg	ggagaccaga	1680
gtagggggag	gggaagagtc	ctgaccagag	gaagacatta	aaaaggtagt	ggggtcgact	1740
agatgaagga	gagcctttct	ctctgggcaa	gagcggtgca	atggtgtgta	aaggtagctg	1800
agaagacgaa	aagggcaagc	atcttcctgc	taccaggctg	gggaggccca	ggcccacgac	1860
cccgaggaga	gggaacgcag	ggagactgag	gtgacccttc	tttcccccg	ggcccggtcg	1920
tgtgtttcgg	tgtctctttt	ctgttgagcc	cttaccttga	cccaggcgct	gccggggcct	1980
gggcccgggc	tgcgggcgac	ggcactcccg	ggaggcagcg	agactcgagt	taggcccac	2040
gcggcgccac	ggcgtttctc	ggccgggaat	ggcccgtagc	cgtgaggtgg	gggtgggggg	2100
cagaaaaggc	ggagcgagcc	cgagcgggga	gggggagggc	caggggcgga	ggggccgggc	2160
actactgtgt	tggcggactg	gcgggactag	ggctgcgtga	gtctctgagc	gcaggcgggc	2220
ggcgcccgcc	ctcccccg	cgcgcgagcg	ggcgagcg	gcggcagctc	actcagcccg	2280
ctgcccagc	ggaaacgcca	ctgaccgcac	ggggattccc	agtgcggcg	ccaggggcac	2340
gcgggacacg	ccccctcccg	ccgcgccatt	ggcctctccg	cccaccgccc	cacacttatt	2400
ggccggtgcg	ccgccaatca	gcggaggctg	ccggggccgc	ctaaagaaga	ggctgtgctt	2460
tggggctccg	gctcctcaga	gagcctcggc	taggtagggg	atcgggactc	tggcgggagg	2520
gcggcttggt	gcgtttgceg	ggatgggcgg	ccgcggcagg	ccctccgagc	gtggtggagc	2580
cgttctgtga	gcagccgggg	tacgagtcgt	gagcctggaa	ggggcaagcg	ggtggtgggc	2640
aggaatgcgg	tccgccctgc	agcaaccgga	gggggaggga	gaagggagcg	gaaaagtctc	2700
caccggacgc	ggccatggct	cggggggggg	ggggcagcgg	aggagcgctt	ccggccgacg	2760
tctcgtcgct	gattggcttc	ttttcctccc	gccgtgtgtg	aaaacacaaa	tggcgtgttt	2820
tggttggcgt	aaggcgctg	tcagttaacg	gcagccggag	tgccgagccg	ccggcagcct	2880
cgctctgccc	actgggtggg	gcgggaggta	ggtggggtag	ggcgagctgg	acgtgcgggc	2940
gcggtcgggc	tctggcgggg	cgggggaggg	gagggagggt	cagcgaaagt	agctcgcgcg	3000
cgagcgcccg	cccaccctcc	ccttcctctg	ggggagtcgt	tttaccgccc	gccggccggg	3060
cctcgtcgct	tgattggctc	tgggggcccc	gaaaactggc	ccttgccatt	ggctcgtgtt	3120
cgtgcaagtt	gagtcacatc	gccggccagc	ggggggcgcg	aggaggcgct	cccaggttcc	3180
ggccctcccc	ctggccccgc	gccgcagagt	ctggcccgcg	gcccctgcgc	aacgtggcac	3240
gaagcgcgcg	ctggggggcg	ggacgggcag	tagggctgag	cggtcgcggg	gcgggtgcaa	3300
gcacgtttcc	gacttgagtt	gcctcaagag	gggcgtgctg	agccagacct	ccatcgcgca	3360
ctccggggag	tggagggaag	gagcgagggc	tcagttgggc	tgttttgag	gcagggaagca	3420
cttgctctcc	caaagtgcgt	ctgagttggt	atcagtaagg	gagctgcagt	ggagtaggcg	3480
gggagaaggc	cgcacccctc	tccggagggg	ggaggggagt	gttgcaatac	ctttctggga	3540
gttctctgct	gcctcctggc	ttctgaggac	cgccctgggc	ctggggagaat	cccttccccc	3600
tcttcctcgt	tgatctgcaa	ctccagtcct	tctagaagat	gggcgggagt	cttctgggca	3660
ggcttaaaag	ctaacctggt	gtgtgggcgt	tgtcctgcag	gggaattgaa	caggtgtaaa	3720
attggaggga	caagacttcc	cacagatttt	cggttttgtc	gggaagtttt	ttaatagggg	3780
caaataagga	aaatgggagg	ataggtagtc	atctgggggt	ttatgcagca	aaactacagg	3840
ttattattgc	tttgtatccg	cctcgagata	ttttccatcg	aggtagatta	aagacatgct	3900
cacccgagtt	ttatactctc	ctgcttgaga	tccttactac	agtatgaaat	tacagtgtcg	3960
cgagttagac	tatgtaagca	gaattttaat	cattttttaa	gagccagta	cttcataatc	4020
atttctcccg	ctccttctgc	agccttatca	aaaggatttt	tagaacactc	atttttagccc	4080
cattttcatt	tattatactg	gcttatccaa	cccctagaca	gagcattggc	attttccctt	4140
tcttgatctt	agaagtctga	tgactcatga	aaccagacag	attagttaca	tacaccacaa	4200
atcgaggctg	tagctggggc	ctcaacactg	cagttctttt	ataactcctt	agtacacttt	4260
ttgttgatcc	tttgccctga	tccttaattt	tcagtgctta	tcacctctcc	cgtcagtggt	4320
gttccacatt	tgggcctatt	ctcagtcag	ggagttttac	aacaatagat	gtattgagaa	4380
tccaacctaa	agcttaactt	tccactccca	tgaatgcctc	tctccttttt	ctccatttat	4440
aaactgagct	attaaccatt	aatggttcca	ggtggatgtc	tcctcccat	attacctgat	4500
gtatcttaca	tattgccagg	ctgatatttt	aagacattaa	aaggtatatt	tcattattga	4560
gccacatggt	attgattact	gcttactaaa	attttgtcat	tgtacacatc	tgtaaaagggt	4620

ggttccctttt	ggaatgcaaa	gttcaggtgt	ttgttgtctt	tcctgacctt	aggtcttgtg	4680
agcttgattt	ttttctattt	aagcagtgct	ttctcttgga	ctggcttgac	tcatggcatt	4740
ctacacgtta	ttgctggctt	aaatgtgatt	ttgccaagct	tcttcaggac	ctataatttt	4800
gcttgacttg	tagccaaaca	caagtaaaat	gattaagcaa	caaatgtatt	tgtgaagctt	4860
ggtttttagg	ttgttgtgtt	gtgtgtgctt	gtgctctata	ataatactat	ccaggggctg	4920
gagaggtggc	tcggagttca	agagcacaga	ctgctcttcc	agaagtcttg	agttcaattc	4980
ccagcaacca	catggtggct	cacaaccatc	tgtaatggga	tctgatgccc	tcttctgggtg	5040
tgtctgaaga	ccacaagtgt	attcacatta	aataaataaa	tcctccttct	tcttcttttt	5100
ttttttttta	aagagaatac	tgtctccagt	agaatttact	gaagtaatga	aatactttgt	5160
gtttgttcca	atatggtagc	caataatcaa	attactcttt	aagcactgga	aatgttacca	5220
aggaactaat	ttttatttga	agtgttaact	tggacagagg	agccataact	gcagacttgt	5280
gggatacaga	agaccaatgc	agactttta	gtcttttctc	ttacactaag	caataaagaa	5340
ataaaaaattg	aacttctagt	atcctatttg	tttaaactgc	tagctttact	taacttttgt	5400
gcttcatcta	tacaaagctg	aaagctaagt	ctgcagccat	tactaaacat	gaaagcaagt	5460
aatgataatt	ttggatttca	aaaatgtagg	gccagagttt	agccagccag	tgggtggtgct	5520
tgcccttatg	cctttaatcc	cagcactctg	gaggcagaga	caggcagatc	tctgagtttg	5580
agcccagcct	ggctctacaca	tcaagttcta	tctaggatag	ccaggaatac	acacagaaac	5640
cctgttgggg	aggggggctc	tgagatttca	taaaattata	attgaagcat	tcctaatga	5700
gccactatgg	atgtggctaa	atccgtctac	ctttctgatg	agatttggtt	attatttttt	5760
ctgtctctgc	tggttggttg	gtcttttgac	actgtgggct	ttctttaaag	cctccttctc	5820
gccactgtgt	ctcttggttg	ctactaactt	ccgatggctt	aaatggcatg	gctttttgct	5880
ttctaagggc	agctgctgag	atgtgcagcc	tgatttccag	gggtgggttg	ggaaattctt	5940
caaacactaa	aatgtctcct	taattttttt	tttaaaaaat	gggttatata	ataaacctca	6000
taaaatagtt	atgaggagtg	aggtggacta	atattaaatg	agtcctccc	ctataaaaga	6060
gctattaagg	ctttttgtct	tatacttaac	ttttttttta	aatgtggtat	ctttagaacc	6120
aagggtctta	gagttttagt	atacagaaac	tgttgcatcg	cttaatcaga	ttttctagtt	6180
tcaaatccag	agaattccaa	ttcttcacag	ccaaagtcaa	attaagaatt	tctgactttt	6240
aatgttaatt	tgcttactgt	gaatataaaa	atgtagctt	ttcctgaggc	agggctcac	6300
tatgtatctc	tgctgatct	gcaacaagat	atgtagacta	aagttctgcc	tgctttgtc	6360
tcctgaatac	taaggttaaa	atgtagtaat	acttttgga	cttgagggc	agattctttt	6420
ataggggaca	cactaaggga	gcttggttga	tagtttgtaa	aatgtgtttc	aagtgtatga	6480
aacttgaatt	attatcaccg	caacctactt	tttaaaaaaa	aaagccaggc	ctgttagagc	6540
atgcttaagg	gactccctag	acttgctgag	cacacaagag	tagttacttg	gcaggctcct	6600
ggtgagagca	tatttcaaaa	aacaaggcag	acaaccaaga	aactacagtt	aaggttacct	6660
gtcttttaaac	catctgcata	tacacaggga	tattaaaata	ttccaaataa	tatttcaatt	6720
aagttttccc	ccatcaaatt	gggacatgga	tttctccggt	gaataggcag	agttggaaac	6780
taaacaaatg	ttggttttgt	gattttgtga	attgttttca	agtgatagtt	aaagcccatg	6840
agatacagaa	caaagctgct	atcttggttt	ctatttggtt	atactcagaa	gcacttcttt	6900
gggtttccct	gcactatcct	gatcatgtgc	taggcctacc	ttaggctgat	tggtgttcaa	6960
ataaacttaa	gtttcctgtc	aggtgatgtc	atatgatttc	atatatcaag	gcaaaacatg	7020
ttatatatgt	taaacatttg	tacttaattg	gaaagttagg	tctttgtggg	tttgattttt	7080
aattttcaaa	acctgagcta	aataagtcac	ttttacatgt	cttacatttg	gtggaattgt	7140
ataattgtgg	tttgcaggca	agactctctg	acctagtaac	cctacctata	gagcactttg	7200
ctgggtcaca	agtctaggag	tcangcattt	caccttgaag	ttgagacgtt	ttgttagtgt	7260
atactagttt	atatgtttga	ggacatgttt	atccagaaga	tattcaggac	tatttttgac	7320
tgggctaagg	aattgattct	gattagcact	gttagtgagc	attgagtggc	ctttaggctt	7380
gaattggagt	cacttgtata	tctcaaataa	tgctggcctt	ttttaaaaag	cccttgttct	7440
ttatcacect	gtttttctaca	taatttttgt	tcaaagaaat	acttgttttg	atctcctttt	7500
gacaacaata	gcatgttttc	aagccatatt	tttttttctt	tttttttttt	tttttttctt	7560
ttcgagacag	ggtttctctg	tatagccctg	gctgtcctgg	aactcacttt	gtagaccagg	7620
ctggcctcga	actcagaaat	ccgcctgect	ctgcctcctg	agtgccggga	ttaaaggcgt	7680
gcaccaccac	gcctggctaa	gttggtatatt	ttgttatata	actataacca	atactaactc	7740
cactgggtgg	atttttaatt	cagtcagtag	tcttaagtgg	tctttatttg	cccttcatta	7800
aaatctactg	ttcactctaa	cagaggctgt	tggtagtagt	ggcacttaag	caacttcta	7860
cggatatact	agcagattaa	gggtcaggga	tagaaactag	tctagcgttt	tgtataccta	7920
ccagctttat	actaccttgt	tctgatagaa	atatttcagg	acatctagag	tgtactataa	7980
ggttgatggg	aagcttataa	ggaacttgaa	agtggagtaa	ctactccatt	tctctgaggg	8040
gagaattaaa	atttttgacc	aagtgtttgt	gagccactga	gaatggtctc	agaacataac	8100
ttcttaagga	accttcccag	attgcccctca	acactgcacc	acatttggtc	ctgcttgaac	8160
attgccatgg	ctcttaaaag	cttaatttaag	aatatttaatt	tgtaatttat	tggttttctc	8220
ccttttagatc	attccttgag	gacaggacag	tgcttgttta	aggctatatt	tctgctgtct	8280
gagcagcaac	aggtcttcga	gatcaacatg	atgttcataa	tcccagatg	ttgccattta	8340
tgttctcaga	agcaagcaga	ggcatgatgg	tcagtgcacg	taatgtcact	gtgttaaatg	8400
ttgctatgca	gtttggattt	ttctaattga	gtgtaggtag	aacatatgtg	ttctgtatga	8460
attaactctt	taagttacac	cttgataaat	ccatgcaatg	tggtatgcaa	ttaccatttt	8520
aagtattgta	gttttctttg	tatgtgagga	taaaggtgtt	tgtcataaaa	tggtttgaac	8580
atttccccaa	agttccaaat	tataaaaacca	caacgttaga	acttatttat	gaacaatggt	8640

tgtagtttca	tgctttttaa	atgcttaatt	attcaattaa	caccgtttgt	gttataatat	8700
atataaaact	gacatgtaga	agtgtttgtc	cagaacattt	cttaaatgta	tactgtcttt	8760
agagagttta	atatagcatg	tcttttgcaa	catactaact	tttgtgttgg	tgcgagcaat	8820
atttgttagt	cattttgaaa	ggagtcattt	caatgagtgt	cagattgttt	tgaatgttat	8880
tgaacatttt	aaatgcagac	ttgttcgtgt	tttagaaagc	aaaactgtca	gaagctttga	8940
actagaaatt	aaaaagctga	agtatttcag	aagggaata	agctacttgc	tgtattagtt	9000
gaaggaaaagt	gtaataagctt	agaaaattta	aaaccatata	gttgtcattg	ctgaatatct	9060
ggcagatgaa	aagaaatact	cagtggttct	tttgagcaat	ataacagctt	gttatattaa	9120
aaattttccc	cacagatata	aactctaata	tataactcat	aaatgttaca	aatggatgaa	9180
gcttacaaat	gtggccttgac	ttgtcactgt	gcttgtttta	gttatgtgaa	agtttgcaa	9240
taaaactatg	tcctaaatag	tcaaactgtg	gaatgacttt	ttaactctatt	ggtttgtcta	9300
gaacagttat	gttgccattt	gccctaattg	tgaaagaaaa	agtggggagt	gccttgccac	9360
tggttcattt	tggtgtgaac	caaagagggg	ggcatgcact	tacacttcaa	acatcctttt	9420
gaaagactga	caagtttggg	tcttcacagt	tggaattggg	catccctttt	gtcagggagg	9480
gagggaggga	gggaggctgg	cttgttatgc	tgacaagtgt	gattaaattc	aaactttgag	9540
gtaagttgga	ggaacttgta	cattgttagg	agtgtgacaa	tttggactct	taatgatttg	9600
gtcatacaaa	atgaacctag	accaacttct	ggaagatgta	tataataact	ccatgttaca	9660
ttgatttcac	ctgactaata	cttatccctt	atcaattaaa	tacagaagat	gccagccatc	9720
tgggcctttt	aaccagaaa	tttagtttca	aactcctagg	ttagtgttct	cactgagcta	9780
catcctgatc	tagtcctgaa	aataggacca	ccatcacccc	caaaaaaatc	tcaataaga	9840
tttatgctag	tgtttcaaaa	ttttaggaat	aggtaagatt	agaaagtttt	aaattttgag	9900
aaatggcttc	tctagaaa	gtcatagat	gaacactgaa	tggtccttaa	agagctaga	9960
aaactggtag	tgagcacaca	ggactgagag	gtctttcttg	aaaagcatgt	attgctttac	10020
gtgggtcaca	gaaggcaggc	aggaagaact	tggtgtgaaa	ctggtgtctt	aagtggctaa	10080
catcttcaca	actgatgagc	aagaacttta	tcctgatgca	aaaaccatcc	aaacaaacta	10140
agtgaaggt	ggcaatggat	cccaggctgc	tctagaggag	gacttgactt	ctcatcccat	10200
caccacacac	agatagctca	tagactgcca	attaacacca	gcttctagcc	tccacaggca	10260
cctgcactgg	tacacataat	ttcacacaaa	cacagtaaga	agccttccac	ctggcatggg	10320
attgcttatc	tttagttccc	aacacttggg	aggcagaggc	cagccagggc	tatgtgacaa	10380
aaaccttgct	tagaggagaa	acttcatagc	ttatttccta	ttcacgtaac	caggttagca	10440
aaattttacca	gccagagatg	aagctaacag	tgtccactat	atttgtagt	ttttaagtca	10500
atttttttaa	tatacttaat	agaattaaag	ctatggtgaa	ccaagtacaa	acctggtgta	10560
ttaaccttag	aacttagcat	aaaaagtagt	tcatttggct	agtaaatatt	aaatgcttac	10620
tggaagat	tatgtcagga	acttggtaaa	tggtgatgaa	acaatcatag	ttgtacatct	10680
tggttctgtg	atcaccttgg	tttgaggtaa	aagtgggtcc	tttgatcaag	gatggaattt	10740
taagtttata	ttcaatcaat	aatgtattat	tttgtgattg	caaaattgcc	tatctagggt	10800
ataaaacctt	taaaaatttc	ataataccag	ttctactctc	agttactaat	tccaaaaagc	10860
cactgactat	ggtgccaatg	tggttctgt	tctcaaagga	aggattgtct	gtgcccttta	10920
ttctaataga	atacatcacac	tgaaaatcta	agctgaaaga	agccagactt	tcctaaataa	10980
ataactttcc	ataaagctca	aacaaggatt	acttttagga	ggcactgtta	aggaactgat	11040
aagtaatgag	gttacttata	taatgatagt	cccacaagac	tatctgagga	aaaatcagta	11100
caactcgaaa	acagaacaac	cagctaggca	ggaataacag	ggctcccaag	tcaggagggtc	11160
tatccaacac	ccctttctgt	tgagggcccc	agacctacat	attgtataca	aacagggagg	11220
tggtgtgatt	taactctcct	gaggtactct	ggtaaatctt	tgctctgagt	aagcagtaca	11280
gtgtacagtt	tacattttca	tttaaagata	cattagctcc	ctctaccccc	taagactgac	11340
aggcactttg	gggttgggga	gggctttgga	aaataacgct	tccatacact	aaaagagaaa	11400
tttctttaat	taggcttgtt	ggttccatac	atctactggg	gtttctacta	cttagtaata	11460
ttataatagt	cacacaagca	tctttgtctt	gttttaggtg	tatatattat	ttaaggcaga	11520
tgataaaaact	gtagatctta	agggatgctt	ctgcttctga	gatgatacaa	agaatttaga	11580
ccataaaaaca	gtagggttga	caagcaatag	aatatggcct	aaagtgttct	gacacttaga	11640
agccaagcag	tgtaggcttc	ttaagaaata	ccattacaat	caccttgcta	gaaatcaagc	11700
attctggagt	ggtcaagcag	tgtaacctgt	actgtaagtt	acttttctgc	tatttttctc	11760
ccaaagcaag	ttctttatgc	tgatatttcc	agtgttagga	actacaaata	ttaataagtt	11820
gtcttctact	ttttctttac	caaggagggt	ctcttctctc	atcttgatct	gaaggatgaa	11880
caaaggcttg	agcagtgcgc	tttagaagat	aaactgcagc	atgaaggccc	ccgatgttca	11940
cccagactac	atggaccttt	cgccacacat	gtccatttcc	agataaggcc	tgccacacac	12000
aaaaaacata	agtcattagg	ctaccagtct	gattctaaaa	caacctaaaa	tcttcccatc	12060
taaatgctat	gggtggtggg	ttggaaagtt	gactcagaaa	atcacttgct	gttttttagag	12120
aggactctggg	tctcagttct	gatacattgt	ggcttacaac	tataactcca	gttctagggg	12180
gtccatccaa	catcctcttc	tggtgagggc	accaaataaa	tgtattgtgt	acaaaacagg	12240
aggtgagtga	tttaactctc	gtgtatagta	ccttggtaaa	acatttcttg	tcctgagtaa	12300
gcagtacagc	tctgcctgtc	cctgggtctac	agacacggct	catttcccgga	aggcaagctg	12360
gatagagatt	ccaatttctc	ttcttggatc	ccatcctata	aaagaaggtc	aagttaatc	12420
tattgcaaaa	ggtaaatagg	tagtttctta	catgagacaa	gaacaaatct	taggtgtgaa	12480
gcagtcactc	tttacaggcc	agagcctcta	ttctatgcca	atgaaggaaa	ctgttagtcc	12540
agtggttatg	cgttatgcca	gtgtatgcca	acactttttt	acactttttt	tttaaacac	12600
tgcaacttag	cttattgaag	acaaaccacg	agtagaaatc	tgtccaagaa	gcaagtgtct	12660

ctcagcctac	aatgtggaat	aggaccatgt	aatggtacag	tgagtgaat	gaattatggc	12720
atgtttttct	gactgagaag	acagtacaat	aaaaggtaaa	ctcatggtat	ttatttataa	12780
agaatccaat	ttctaccttt	ttccaaatgg	catactgtt	acaataatat	ccacagaagc	12840
agttctcagt	gggaggttgc	agatatccca	ctgaacagca	tcaatgggca	aaccccaggt	12900
tgtttttctg	tggagacaaa	ggtaagatat	ttcaatatat	tttcccaagc	taatgagatg	12960
gctcagcaaa	taatggtact	ggccattaag	tctcatgacc	tgagcttgat	cctcagggac	13020
catgtgttac	aaggagagac	ctaaatcctt	cagttggact	tcaatcttct	accctcatgt	13080
ccacacacaa	ataaatacaa	taaaaaacat	tctgcagtcg	aatttctaaa	agggcgaat	13139

<210> 4

<211> 1073

<212> DNA

<213> Mus musculus

<220>

<223> Description: 5' arm of Rosa 26

<400> 4

caggccctcc	gagcgtggtg	gagccgttct	gtgagacagc	cgggtacgag	tctgacgct	60
ggaaggggca	agcgggtggt	gggcaggaat	gcggtccgcc	ctgcagcaac	cggaggggga	120
gggagaagg	agcggaaaag	tctccaccgg	acgcggccat	ggctcggggg	ggggggggca	180
gcggaggagc	gcttccggcc	gacgtctcgt	cgctgattgg	cttcttttcc	tcccgcctgt	240
tgtgaaaaca	caaatggcgt	gttttggttg	gcgtaaggcg	cctgtcagtt	aacggcagcc	300
ggagtgcgca	gccgcgggca	gcctcgtctc	gccactggg	tggggcgggg	ggtaggtggg	360
gtgaggcgag	ctggacgtgc	gggcgcggtc	ggcctctggc	ggggcggggg	aggggagggg	420
gggtcagcga	aagtagctcg	cgcgcgagcg	gccgccacc	ctcccttcc	tctgggggag	480
tctgtttacc	gcgcgcggc	cgggcctcgt	cgtctgattg	gctctcgggg	cccagaaaac	540
tggcccttgc	cattggctcg	tgttcgtgca	agttgagtc	atccgcgggc	cagcgggggc	600
ggcgaggagg	cgctcccagg	ttccggccct	ccctcgggcc	ccgcgcggca	gagctctggc	660
gcgcgccccct	gogcaacgtg	gcaggaagcg	cgcgctgggg	gcggggacgg	gcagtagggc	720
tgagcggctg	cggggcgggg	gcaagcacgt	ttccgacttg	agttgcctca	agagggcggt	780
gctgagccag	acctccatcg	cgcactccgg	ggagtggagg	gaaggagcga	gggctcagtt	840
gggctgtttt	ggaggcagga	agcaattgct	ctcccaaagt	cgctctgagt	tggtatcagt	900
aaggggagctg	cagtggagta	ggcggggaga	aggccgcacc	cttctccgga	ggggggaggg	960
gagtgttgca	atacctttct	gggagttctc	tgctgcctcc	tggcttctga	ggaccgcctt	1020
gggcctggga	gaatcccttc	cccctcttcc	ctcgtgatct	gcaactccag	tct	1073

<210> 5

<211> 4333

<212> DNA

<213> Mus musculus

<220>

<223> Description: 3' arm of Rosa 26

<400> 5

tagaagatgg	gcgggagttc	tctgggcagg	cttaaaggct	aacctgggtg	gtgggcgttg	60
tctgcaggg	gaattgaaca	ggtgtaaaat	tggagggaca	agacttccca	cagattttcg	120
gttttgcgg	gaagtttttt	aataggggca	aataaggaaa	atgggaggat	aggtagtcat	180
ctgggggttt	atgcagcaaa	actacagggt	attattgctt	gtgatccgcc	tccgagattt	240
ttocatcgag	gtagattaaa	gacatgctca	cccaggtttt	atactctcct	gcttgagatc	300
cttactacag	tatgaaatta	cagtgtcgcg	agttagacta	tgtaagcaga	attttaatca	360
tttttaaga	gcccagtaact	tcatatccat	ttctcccgct	ccttctgcag	ccttatcaaa	420
aggatatttt	gaacactcat	tttagcccca	ttttcattta	ttatactggc	ttatccaacc	480
cctagacaga	gcattggcat	tttccctttc	ctgatcttag	aagtctgatg	actcatgaaa	540
ccagacagat	tagttacata	caccacaaat	cgaggctgta	gctggggcct	caacactgca	600
gttcttttat	aactccttag	tacacttttt	gttgatcctt	tgcttgatc	cttaattttc	660
agtgtctatc	acctctcccg	tcagtgggtg	tccacatttg	ggcctattct	cagtcagggg	720
agtttttaaa	caatagatgt	attgagaatc	caacctaaag	cttaactttc	cactcccatg	780
aatgcctctc	tcctttttct	ccattttata	actgagctat	taaccattaa	tggttccagg	840
tggatgtctc	ctcccatat	tacctgatgt	atcttacata	ttgccaggct	gatattttta	900
gacattaaaa	ggtatatatt	attattgagc	cacatgggtat	tgattactgc	ttactaaaa	960
tttgtcattg	tacacatctg	taaaagggtg	ttccttttgg	aatgcaaagt	tcagggtgtt	1020
gttgtctttc	ctgacctaa	gtcttgtgag	cttgtatttt	ttctatttaa	gcagtgcttt	1080
ctcttgagct	ggcttgactc	atggcattct	acacgttatt	gctgggtctaa	atgtgatttt	1140

```

gccaaagcttc ttcaggacct ataattttgc ttgacttgta gccaaacaca agtaaaatga 1200
ttaagcaaca aatgtatttg tgaagcttgg tttttagggt gttgtgttgt gtgtgcttgt 1260
gctctataat aatactatcc aggggctgga gaggtggctc ggagttcaag agcacagact 1320
gctcttcagc aagtcctgag ttcaattccc agcaaccaca tgggtggctca caaccatctg 1380
taatgggata tgatgccttc ttctgggtgt tctgaagacc acaagtgtat tcacattaaa 1440
taaataaatc ctccctcttc ttcttttttt tttttttaaa gagaatactg tctccagtag 1500
aatttactga agtaatgaaa tactttgtgt ttgttccaat atggtagcca ataatacaat 1560
tactctttaa gcactggaaa tgttaccaag gaactaattt ttatttgaag tgtaactgtg 1620
gacagaggag ccataactgc agacttgtgg gatacagaag accaatgcag actttaatgt 1680
cttttctctt acactaagca ataaagaaat aaaaattgaa cttctagtat cctatttgtt 1740
taaactgcta gctttactta acttttgtgc ttcacttata caaagctgaa agctaagctc 1800
gcagccatta ctaaacatga aagcaagtaa tgataatttt ggatttcaaa aatgtagggc 1860
cagagtttag ccagccagtg gtggtgcttg cctttatgcc tttaatcca gcactctgga 1920
ggcagagaca ggcagatctc tgagtttgag ccagcctgg tctacacatc aagtctctatc 1980
taggatagcc aggaatacac acagaaaccc tgttggggag gggggctctg agatttcata 2040
aaattataat tgaagcattc cctaattgag cactatggat gtggctaaat cctgtcactc 2100
ttctgatgag atttgggtat tttttttct gtctctgctg ttggttgggt ctttgacac 2160
tgtgggcttt ctttaaagcc tcttctctgc catgtggtct cttgtttgct actaacttcc 2220
catggcttaa atggcatggc tttttgcctt ctaagggcag ctgctgagat ttgcagcctg 2280
atttccaggg tggggttggg aaatctttca aacactaaaa ttgtccttta atttttttt 2340
taaaaaatgg gttatataat aaacctcata aaatagttat gaggagttag gtggaactat 2400
attaaatgag tccctccctc ataaagagc tattaaggct ttttgtctta tacttaactt 2460
ttttttttaa tgtggtatct ttagaaccaa gggctcttaga gtttttagtat acagaaactg 2520
ttgcatcgct taatcagatt ttctagtttc aaatccagag aatccaaatt cttcacagcc 2580
aaagtcaaat taagaatttc tgacttttaa tgttaatttg cttactgtga atataaaaaat 2640
gatagctttt cctgaggcag ggtctcacta tgtatctctg cctgatctgc aacaagatat 2700
gtagactaaa gttctgcctg cttttgtctc ctgaatacta aggttaaaat gtagtatac 2760
ttttggaact tgcaggtcag attcttttat aggggacaca ctaggggagc ttgggtgata 2820
gttggtaaaa tgtgtttcaa gtgatgaaaa cttgaattat tatcaccgca acctactttt 2880
taaaaaaaaa agccaggcct gtttagagcat gcttaaggga tccctaggac ttgctgagca 2940
cacaagagta gttacttggc aggtcctctg tgagagcata tttcaaaaaa caaggcagac 3000
aaccaagaaa ctacagttaa ggttacctgt ctttaaacca tctgcatata cacagggata 3060
ttaaaatatt ccaaataata tttcattcaa gttttcccc atcaaatggg gacatggatt 3120
tctccggtga ataggcagag ttggaaacta aacaaatgtt ggttttgtga tttgtgaaat 3180
tgttttcaag tgatagttaa agcccatgag atacagaaca aagctgctat ttcgaggtct 3240
cttggtttat actcagaagc acttcttttg gtttccctgc actatcctga tcatgtgcta 3300
ggcctacctt aggttgattg ttgttcaaat aaacttaagt ttcctgtcag gtgatgtcat 3360
atgatttcac atatcaaggc aaaacatggt atatatgtta aacatttgta cttaattgtg 3420
aagtttaggtc tttgtgggtt tgatttttaa ttttcaaaac ctgagctaaa taagtcattt 3480
ttacatgtct tacatttggt ggaattgtat aattgtgggt tgcaggcaag actctctgac 3540
ctagtaaccc tacctataga gcactttgct gggtcacaag tctaggagtc aagcatttca 3600
ccttgaagtt gagacgtttt gttagtgtat actagtttat atgttggagg acatgtttat 3660
ccagaagata ttcaggacta tttttagctg ggctaaggaa ttgattctga ttagcactgt 3720
tagtgagcat tgagtggcct ttaggcttga attggagtca cttgtatatc tcaaataatg 3780
ctggcctttt ttaaaaagcc cttgttcttt atcacctgtt tttctacata atttttgttc 3840
aaagaaatac ttgtttggat ctoccttttg caacaatagc atgttttcaa gccatatttt 3900
ttttcctttt tttttttttt tttgggtttt cgagacaggg tttctctgta tagccttggc 3960
tgtcttgtaa ctactttgt agaccaggct ggctcgaac tcagaaatcc gcctgcctc 4020
gctcctgag tgccgggatt aaaggcgtgc accaccagc ctggctaagt tggatatttt 4080
gttatataac tataaccaat actaactcca ctgggtggat ttttaattca gtcagtagtc 4140
ttaagtggtc tttattggcc cttcattaaa atctactgtt cactctaaca gaggctgttg 4200
gtactagtgg cacttaagca acttctctac gatatactag cagattaagg gtcagggata 4260
gaaactagtc tagcgttttg tatacctacc agctttatac taccttgttc tgatagaaat 4320
atttcaggac atc 4333

```

<210> 6

<211> 13544

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Rosa 26
targeting vector pOMP3 with zsgreen

<400> 6

ctagggataa cagggttaata tagccgcggc aggcctctccg agcgtgggtg agccgttctg 60

tgagacagcc	gggtacgagt	cgtgacgctg	gaaggggcaa	gcgggtggtg	ggcaggaatg	120
cgggtccgccc	tgcagcaacc	ggagggggag	ggagaaggga	gcggaaaagt	ctccaccgga	180
cgcgcccatg	gctcgggggg	gggggggag	cggaggagcg	cttccggccg	acgtctcgtc	240
gctgattggc	ttcttttcct	cccgcggtgt	gtgaaaacac	aaatggcggtg	ttttggttgg	300
cgtaaggcgc	ctgtcagtta	acggcagccg	gagtgcgcag	ccgcggcgag	cctcgtctctg	360
cccactgggt	ggggcgggag	gtaggtgggg	tgaggcgagc	tggacgtgcg	ggcgcggtcg	420
gcctctggcg	gggcggggga	ggggaggag	ggtcagcgaa	agtagctcgc	gcgcgagcgg	480
ccgcccaccc	tcoccttcct	ctgggggagt	cgttttacc	gccgcccggc	gggcctcgtc	540
gtctgattgg	ctctcggggc	ccagaaaact	ggccttgcc	attggctcgt	gttcgtgcaa	600
gttgagtcca	tccgcgggcc	agcggggggc	gcgaggaggg	gctcccagg	tccggccctc	660
ccctcgcccc	cgcgcgcgag	agtctggccg	cgcgcccctg	cgcaacgtgg	caggaagcgc	720
gcgctggggg	cggggacggg	cagtagggct	gagcggctgc	ggggcggtg	caagcacgtt	780
tcgcacttga	gttgccctcaa	gagggcggtg	ctgagccaga	cctccatcgc	gcactccggg	840
gagtgagggg	aaggagcgag	ggctcagttg	ggctgttttg	gaggcaggaa	gcacttgctc	900
toccaaagtc	gctctgagtt	gttatcagta	agggagctgc	agtggagtag	gcggggagaa	960
ggcgcacccc	ttctccggag	gggggagggg	agtgttgcaa	tacctttctg	ggagttctct	1020
gctgcctcct	ggcttctgag	gaccgcctcg	ggcctgggag	aatcccttcc	ccctcttccc	1080
tcgtgatctg	caactccagt	ctttctaggt	aaccgatata	cctgcagggg	tgacctgcac	1140
gtctagggcg	cagtagtcca	gggtttcctt	gatgatgtca	tacttatcct	gtcccttttt	1200
tttccacagc	tgcgggttga	ggacaaactc	ttcgcggtct	ttccagta	cctgcagggtg	1260
actgactgag	tgcactaggg	gtaaccataa	cttcgtataa	tgatgctat	acgaagtatt	1320
ttataatcta	gaactagtgg	atccacgatt	cgaggggccc	tgacaggtcaa	ttctaccggg	1380
taggggaggg	cctttttcca	aggcagctcg	gagcatgcgc	tttagcagcc	ccgctgggca	1440
cttggcgcta	cacaagtggc	ctctggcctc	gcacacattc	cacatccacc	ggtaggcgcc	1500
aacgggctcc	gttcttttgt	ggcccttcg	cgccaccttc	tactcctccc	ctagtcagg	1560
agttcccccc	cgcggcgag	ctcgcgtcgt	gcaggacgtg	acaaatggaa	gtagcacgtc	1620
tcactagtct	cgtgcagatg	gacagcaccg	ctgagcaatg	gaagcggtg	ggcctttggg	1680
gcagcgccca	atagcagctt	tgctccttcg	cttctggggc	tcagaggctg	ggaagggtg	1740
ggtcgggggg	cgggctcagg	gcgggctcag	gggcggggcg	ggcgcccgaa	ggtcctccgg	1800
aggcccggca	ttctgcacgc	ttcaaaagcg	cacgtctgcc	gcgctgttct	cctcttccct	1860
atctccgggc	ctttcgacct	gcagccaata	tgggatcggc	cattgaacaa	gatggattgc	1920
acgcagggtt	tccggccgct	tgggtggaga	ggctattcgg	ctatgactgg	gcacaacaga	1980
caatcgggtc	ctctgatgcc	gccgtgttcc	ggctgtcagc	gcaggggcgc	ccggttcttt	2040
ttgtcaagac	cgacctgtcc	ggtgccctga	atgaactgca	ggacgaggca	gcgcggtat	2100
cgtggctggc	cacgacgggc	gttcccttgc	cagctgtgct	cgacgttgtc	actgaagcgg	2160
gaagggactg	gctgctattg	ggcgaagtgc	cggggcagga	tctcctgtca	tctcaccttg	2220
ctcctgccga	gaaagtatcc	atcatggctg	atgcaatgog	gcggctgcat	acgcttgatc	2280
cggtacacct	cccattcgac	caccaagcga	aacatcgcat	cgagcgagca	cgtactcgga	2340
tggaaagccg	tcttctcgat	caggatgac	tggacgaaga	gcatacggg	ctcgcgccag	2400
ccgaactgtt	cgccaggctc	aaggcgcgca	tgcccgacgg	cgaggatctc	gtcgtgacct	2460
atggcgatgc	ctgcttgccg	aatatcatgg	tggaaaatgg	ccgcttttct	ggattcatcg	2520
actgtggccg	gctgggtgtg	gcggaccgct	atcaggacat	agcgttggtc	acccgtgata	2580
ttgtgaaga	gcttggcggc	gaatgggctg	accgcttcc	cgtgctttac	ggatcgccg	2640
ctcccgattc	gcagcgcatc	gccttctatc	gccttcttga	cgagttcttc	tgaggggatc	2700
gatccgctgt	aagtctgcag	aaattgatga	tctattaaac	aataaagatg	tccactaaaa	2760
tggaaagttt	tcctgtcata	ctttgttaag	aagggtgaga	acagagtacc	tacattttga	2820
atggaaggat	tggagctacg	gggtgggggg	tgggtgggga	ttagataaat	gcctgctcct	2880
tactgaaggc	tctttactat	tgctttatga	taatgtttca	tagttggata	tcataattta	2940
aacaagcaaa	accaaattaa	gggccagctc	attcctccca	ctcatgatct	atagactcat	3000
agatctctcg	tgggatcatt	gtttttctct	tgattccccc	tttgtggttc	taagtactgt	3060
ggtttccaaa	tgtgtcagtt	tcatagcctg	aagaacgaga	tcagcagcct	ctgttccaca	3120
tacacttcat	tctcagtatt	gttttgccaa	gttctaattc	catcagaagc	tgactctaga	3180
tcctgcagga	attcatatga	taacttcgta	taatgtatgc	tatacgaagt	tatggcgcg	3240
cggtaaccca	agttccctata	ctttctagag	aataggaact	tcggaatagg	aacttcaagc	3300
ttagcgcta	gaagatgggc	gggagtcttc	tgggcaggct	taaaaggctaa	cctgggtgtg	3360
ggcgcttgct	ctgcagggga	attgaacagg	tgtaaaatg	gagggacaag	acttcccaca	3420
gattttcggt	tttgtcgga	agttttttaa	taggggcaaa	taaggaaaat	gggaggatag	3480
gtagtcatct	gggtttttat	gcagcaaaac	tacaggttat	tattgcttgt	gatccgcctc	3540
ggagtatttt	ccatcgaggt	agattaaaga	catgctcacc	cgagttttat	actctcctgc	3600
ttgagatcct	tactacagta	tgaattaca	gtgtcgcgag	ttagactatg	taagcagaat	3660
tttaatcatt	tttaagagc	ccagtacttc	atatccattt	ctcccgtcc	ttctgcagcc	3720
ttatcaaaag	gtattttaga	acactcattt	tagccccatt	ttcatttatt	atactggctt	3780
atccaacccc	tagacagagc	attggcattt	tccctttcct	gatcttagaa	gtctgatgac	3840
tcatgaaacc	agacagatta	gttacatata	ccacaaatcg	aggctgtagc	tggggcctca	3900
acactgcagt	tcttttataa	ctccttagta	cactttttgt	tgatcctttg	ccttgatcct	3960
taattttcag	tgtctatcac	ctctcccgct	agtggtgttc	cacatttggg	cctatttcca	4020
gtccagggag	ttttacaaca	atagatgtat	tgagaatcca	acctaagct	taactttcca	4080

ctcccatgaa	tgcctctctc	ctttttctcc	atttataaac	tgagctatta	accattaatg	4140
gttcocagg	gatgtctcct	ccccatatta	cctgatgtat	cttacatatt	gccaggctga	4200
tatftttaaga	cattaaaagg	tatatttcat	tattgagcca	catgggtattg	attactgctt	4260
actaaaat	tgtcattgta	cacatctgta	aaagggtggt	ccttttggaa	tgcaaagttc	4320
agggtgtt	tgtctttcct	gacctaagg	cttgtgagct	tgtattttt	ctattttaagc	4380
agtgccttct	cttggactgg	cttgactcat	ggcattctac	acgttattgc	tggtctaaat	4440
gtgattttgc	caagcttctt	caggacctat	aattttgctt	gacttgtagc	caaacacaag	4500
taaaatgatt	aagcaacaaa	tgtatttgtg	aagcttggtt	tttaggttgt	tgtgttgtgt	4560
gtgcttgtgc	tctataataa	tactatccag	gggctggaga	ggtggctcgg	agttcaagag	4620
cacagactgc	tcttccagaa	gtcctgagtt	caattcccag	caaccacatg	gtggctcaca	4680
accatctgta	atgggatctg	atgcctctt	ctgggtgtgc	tgaagaccac	aagtgtattc	4740
acattaaata	aataaatcct	ccttcttctt	ctttttttt	tttttaaga	gaatactgtc	4800
tcagtagaa	tttactgaag	taatgaaata	ctttgtgtt	gttccaatat	ggtagccaat	4860
aatcaaatta	ctctttaagc	actggaaatg	ttaccaagga	actaattttt	atttgaagt	4920
taactgtgga	cagaggagcc	ataactgcag	acttggtgga	tacagaagac	caatgcagac	4980
tttaattgtct	tttctcttac	actaagcaat	aaagaaataa	aaattgaact	tctagtatcc	5040
tatttgttct	aactgctagc	tttacttaac	tttgtgctt	catctataca	aagctgaaag	5100
ctaagtctgc	agccattact	aaacatgaaa	gcaagtaatg	ataatttttg	atttcaaaaa	5160
tgtagggcca	gagtttagcc	agccagtgg	ggtgcttgcc	tttatgcctt	taatcccagc	5220
actctggagg	cagagacagg	cagatctctg	agtttgagcc	cagcctgggc	tacacatcaa	5280
gttctatcta	ggatagccag	gaatacacac	agaaacctg	ttggggagg	gggctctgag	5340
atttcaataa	attataattg	aagcattccc	taatgagcca	ctatggatgt	ggctaattcc	5400
gtctaccttt	ctgatgagat	ttgggtatta	tttttctgt	ctctgctgtt	gggtgggtct	5460
tttgacactg	tgggctttct	ttaaagcctc	cttctgcca	tgtggtctct	tgtttgtctac	5520
taacttccca	tggcttaaat	ggcatggctt	tttgcttct	aagggcagct	gctgagattt	5580
gcagcctgat	ttccagggtg	gggttgggaa	atctttcaaa	cactaaaatt	gtcctttaat	5640
ttttttttta	aaaaatgggt	tatataataa	acctcataaa	atagttatga	ggagtggagt	5700
ggactaatat	taaatgagtc	cctcccctat	aaaagagcta	tttaaggctt	ttgtcttata	5760
cttaactttt	tttttaaatg	tggatatctt	agaaccaagg	gtcttagagt	tttagtatac	5820
agaaactgtt	gcacgctta	atcagatttt	ctagtttcaa	atccagagaa	tccaaattct	5880
tcacagccaa	agtcaaatta	agaatttctg	acttttaatt	tttaatttgc	tactgtgaat	5940
ataaaaaatga	tagcttttcc	tgaggcagg	tctcactatg	tatctctgcc	tgatctgcaa	6000
caagatatgt	agactaaagt	tctgcctgct	ttgtctctct	gaatactaag	gttaaaatgc	6060
agtaataact	ttggaacttg	cagggtcagat	tcttttatag	gggacacact	aagggtgctt	6120
gggtgatagt	tggtaaaatg	tgtttcaagt	gatgaaaact	tgaattatta	tcaccgcaac	6180
ctacttttta	aaaaaaaaaag	ccaggcctgt	tagagcatgc	tttaaggatc	cctaggactt	6240
gctgagcaca	caagagtagt	tacttggcag	gctcctggtg	agagcatatt	tcaaaaaaca	6300
aggcagacaa	ccaagaaact	acagttaagg	ttacctgtct	ttaaaccatc	tgcatataga	6360
cagggaatatt	aaaatattcc	aaataatatt	tcattcaagt	tttcccccat	caaattggga	6420
catggatttc	tccggtgaat	aggcagagtt	ggaaactaaa	caaagtgttg	ttttgtgatt	6480
tgtgaaattg	ttttcaagt	atagttaaag	cccatgagat	acagaacaaa	gctgctattt	6540
cgaggtctct	tgggtttatac	tcagaagcac	ttctttgggt	ttccctgcac	tatcctgac	6600
atgtgctagg	cctaccttag	gctgattggt	gttcaaataa	acttaagttt	cctgtcaggt	6660
gatgtcatat	gatttcatat	atcaaggcaa	aacatgttat	atatgttaaa	catttgattt	6720
taattgtgaa	gttaggtctt	tgtgggtttg	atttttaatt	ttcaaaacct	gagctaaata	6780
agtcattttt	acatgtctta	catttgggtg	aattgtataa	ttgtggtttg	caggcaagac	6840
tctctgacct	agtaacctta	cctatagagc	actttgctgg	gtcacaaagc	taggagtcac	6900
gcattttcac	ttgaagtiga	gacgttttgt	tagtgtatac	tagtttatat	gttggaggac	6960
atgctttatcc	agaagatatt	caggactatt	tttgactggg	ctaaggaaat	gattctgatt	7020
agcactgtta	gtgagcattg	agtggccttt	aggcttgaat	tggagtcaact	tgtatatctc	7080
aaataatgct	ggcctttttt	aaaaagccct	tgttctttat	caccctgttt	tctacataat	7140
ttttgttcaa	agaaataact	gtttggatct	ccttttgaca	acaatagcat	gttttcaagc	7200
catatttttt	ttcctttttt	ttttttttt	tggtttttcg	agacagggtt	tctctgtata	7260
gccctggctg	tctctggaact	cactttgtag	accaggctgg	cctcgaactc	agaaatccgc	7320
ctgcctcatg	ctcctgagtg	ccgggattaa	aggcgtgcac	caccacgcct	ggctaagttg	7380
gatattttgt	tatataacta	taaccaatac	taactccact	gggtggattt	tttaattcagt	7440
cagtagtctt	aagtggctct	tattggccct	tcattaaaaat	ctactgttca	ctctaacaga	7500
ggctgttgg	actagtggca	cttaagcaac	ttcctacgga	tatactagca	gattaagggt	7560
cagggataga	aactagtcta	gcgttttgta	tacctaccag	ctttatacta	ccttgttctg	7620
atagacacat	tccaggacat	ctagcacggg	caccggtgtt	tttaagatac	attgatgagt	7680
ttggacaaaac	cacaactaga	atgcagtga	aaaaatgctt	tatttgtgaa	atttgtgagt	7740
ctattgcttt	atttgaacc	attataagct	gcaataaaca	agttaacaac	aacaattgca	7800
ttcattttat	gtttcagggt	cagggggagg	tgtgggaggt	tttttaaagc	aagtaaaacc	7860
tctacaaatg	tggatggct	gattatgatc	tagagtgcgc	gccgctcagg	gcaaggcgga	7920
gccggaggcg	atggcgtgct	cggtcagggt	ccacttctgg	ttcttggcgt	cgctgcggtc	7980
ctcgcgggtc	agcttgtgct	ggatgaagtg	ccagtcgggc	atcttgcggg	gcacggactt	8040
ggccttgtac	acggtgtcga	actggcagcg	caagcggcca	ccgtccttca	gcagcaggta	8100

catgctcacg	tcgcccttca	agatgccctg	cttggggcacg	gggatgatct	tctcgcagga	8160
gggctcccag	ttgtcgggtca	tcttcttcat	cacggggccg	tcggcgggga	agttcacgcc	8220
gtagaacttg	gactcgtggt	acatgcagtt	ctcctccacg	ctcacggtga	tgctggcggt	8280
gcagatgcac	acggcgccgt	cctcgaacag	gaaggagcgg	tcccaggtgt	agccggcggg	8340
gcaggagttc	ttgaagtagt	cgacgatgtc	ctgggggtac	tcggtgaaca	cgcggttgcc	8400
gtacatgaag	gcggcggaca	agatgtctc	ggcgaagggc	aagggggccg	cctccaccac	8460
gcacaggttg	atggcctgct	tgcccttgaa	ggggtagccg	atgcctcgc	cggtgatcac	8520
gaacttggtg	ccgtccacgc	agccctccat	gcggtacttc	atggtcatct	ccttggtcag	8580
gccgtgcttg	gactgggcca	tggtggcgac	cggtggatcc	actagtctta	gctgtaggaa	8640
aaagaagaag	gcatgaacat	ggttagcaga	ggctctaggg	ccgccttggg	agtggacacc	8700
tgtggagaga	aaggcaaagt	ggatgtcaat	gtcactcaag	tgataggcca	gatctcaagc	8760
ctgccacacc	tcaagcttga	caacaaaaag	attgtctttt	ctgaccagat	ggacgcggcc	8820
accctcaaag	gcatcacgcg	gggccagggtg	aatatcaa	cctcctcgtt	tttggaact	8880
gacaatctta	gcgcagaagt	catgcccgct	tttgagaggg	agtactcacc	ccaacgctta	8940
ctagagccgc	cggtcacacg	ccagaagccg	aaccccgccc	tgccccgtcc	ccccggaagg	9000
cagccgtccc	cctgcggcag	ccccgaggct	ggagatggag	aaggggacgg	cgggcgcgcg	9060
acgcacgaag	gccctccccg	ccattttcct	tcctgcgggc	gccgcaccgc	ttcgcccgcc	9120
cccgttagag	ggggtgcggc	ggcgctccc	agatttcggc	tcggccagat	ttgggacaaa	9180
ggaagtccct	gcgccctctc	gcacgattac	cataaaaggc	aatggctgcg	gctcgccgcg	9240
cctcgacagc	cgccggcgct	ccggggcgcg	cgcgccctc	ccccgagccc	ttcccggccc	9300
gaggcgggcc	cgccccgccc	ggcaccccca	cctgcgcgca	cccccgccc	ggcacggcga	9360
gccccgcggc	acgccccgca	cggaagcccg	cccccgaagc	cgggccgtgc	tcagcaactc	9420
ggggaggggg	gtgcaggggg	gggttacagc	ccgaccgcg	cgccacacac	ccctgctcac	9480
ccccccacgc	acacaccccg	cacgcagcct	ttgttccct	cgagccccc	ccgcaccgcg	9540
gggcaccgcc	cccgcccgcg	ctccctcgc	gcacacgcgg	agcgacaaa	gccccgcgcc	9600
gcgcccgcag	cgctcacagc	cgccgggcag	cgcgggcgc	acgcggcgct	ccccacgcac	9660
acacacacgc	acgcaccccc	cgagccgctc	cccccgcac	aaaggccct	cccgagccct	9720
tttaaggctt	tcacgcagcc	acagaaaaga	aacgagccgt	cattaaacca	agcgctaatt	9780
acagcccgga	ggagaagggc	cgtcccgcgc	gctcacctgt	gggagtaacg	cggtcagtca	9840
gagccggggc	gggcggcgcg	aggcgggcg	gagcggggca	cgggcggaag	gcaacgcagc	9900
gactcccgc	cgccgcgcgc	ttcgcttttt	atagggcgcg	cgccgcgcgc	gcctcgccat	9960
aaaaggaaac	tttcggagcg	cgccgctctg	attggctgcc	gccgcaactc	tcgcctcgc	10020
cccgcccgcc	ccctcgcccc	gccccgcccc	gcttgccgcg	cgcccccccc	cccccccgcc	10080
ccatcgctgc	acaaaataat	taaaaaataa	ataaatacaa	aattgggggt	ggggaggggg	10140
gggagatggg	gagagtgaag	cagaacgtgg	ggctcacctc	gacccatggg	aatagcgatg	10200
actaatcgt	agatgtactg	ccaagtagga	aagtcccata	aggtcatgta	ctgggcataa	10260
tgccaggcgg	gccatttacc	gtcattgacg	tcaatagggg	gcgtacttgg	catatgatac	10320
acttgatgta	ctgccaagtg	ggcagtttac	cgtaaatagt	ccaccatttg	acgtcaatgg	10380
aaagtcccta	ttggcgttac	tatgggaaca	tacgtcatta	ttgacgtcaa	tgggcggggg	10440
togttggggc	gtcagccagg	cgggccattt	accgtaagtt	atgtaacgcg	gaactccata	10500
tatgggctat	gaactaatga	ccccgtaatt	gattactatt	aataactagt	caataataca	10560
tgtcgaaaac	ctgcaggcat	gcaagctagc	ccaaaacctat	attaccctgc	cctagcgtaa	10620
ctatttaaat	ggccggcccc	tgaggcgccg	ggtggagctc	caattcgccc	tatagttagt	10680
cgtattacra	ttcactggcc	gtcggtttac	aacgtcgtga	ctgggaaaaa	cctggcggtta	10740
cccaacttaa	tcgccttgca	gcacatcccc	ctttcgccag	ctggcgtaat	agcgaagagg	10800
cccgcaccca	tcgcccttcc	caacagttgc	gcagcctgaa	tgccgaatgg	gacgcgcctt	10860
gtagcggcgc	attaagcgcg	gcgggtgtgg	tggttacgcg	cagcgtgacc	gctacacttg	10920
ccagcgccct	agcgcccgct	cctttcgctt	ctttccttcg	ctttctcgcc	acgttcgcgc	10980
gctttccccc	ctaaagctcta	aatcgggggc	tccttttagg	gttccgattt	agtgctttac	11040
ggcacctcga	ccccaaaaaa	cttgattagg	gtgatggttc	acgtagtggg	ccatcgccct	11100
gatagacggt	ttttcgccct	ttgacgttgg	agtcacggtt	ctttaatagt	ggactcttgt	11160
tccaaactgg	aacaacactc	aacccctatct	cgggtctattc	ttttgattta	taagggattt	11220
tgccgatttc	ggcctatttg	ttaaaaaatg	agctgattta	acaaaaattt	aacgcgaatt	11280
ttaacaaaaat	attaacgctt	acaatttagg	tggcactttt	cggggaaatg	tgcgcggaac	11340
ccctatttgt	ttatttttct	aaatacattc	aaatatgtat	ccgctcatga	gacaataaac	11400
ctgataaatg	cttcaataat	attgaaaaag	gaagagtatg	agtattcaac	atttccggtg	11460
cgcccttatt	cccttttttg	cggcattttg	ccttctgtgt	tttgctcacc	cagaaacgct	11520
ggtgaaagta	aaagatgctg	aagatcagtt	gggtgcacga	gtgggttaca	tcgaactgga	11580
tctcaacagc	ggtgaagatcc	ttgagagttt	tcgccccgaa	gaacgttttc	caatgatgag	11640
cactttttaaa	gttctgctat	gtggcgcggt	attatcccggt	attgacgcgg	ggcaagagca	11700
actcggtcgc	cgcatacact	attctcagaa	tgacttggtt	gagtactcac	cagtcacaga	11760
aaagcatctt	acggatggca	tgacagtaag	agaattatgc	agtgtgcga	taaccatgag	11820
tgataacact	gcggccaact	tacttctgac	aacgatcgga	ggaccgaagg	agctaaccgc	11880
ttttttgcac	aacatggggg	atcatgtaac	tcgccttgat	cgttgggaac	cggagctgaa	11940
tgaagccata	ccaaacgacg	agcgtgacac	cacgatgcct	gtagcaatgg	caacaacggt	12000
gcgcaaaacta	ttaaactggc	aactacttac	cttagcttcc	cggcaacaat	taatagactg	12060
gatggaggcg	gataaagttg	caggaccact	tctgcgctcg	gcccttcgcg	ctggtggtt	12120

tattgctgat	aaatctggag	cgggtgagcg	tgggtctcgc	ggatcattg	cagcactggg	12180
gccagatggt	aagccctccc	gtatcgtagt	tatctacacg	acggggagtc	aggcaactat	12240
ggatgaacga	aatagacaga	tcgctgagat	aggtgcctca	ctgattaagc	attggtaact	12300
gtcagaccaa	gtttactcat	atatacttta	gattgattta	aaacttcatt	tttaatttaa	12360
aaggatctag	gtgaagatcc	tttttgataa	tctcatgacc	aaaatccctt	aacgtgagtt	12420
ttcgttccac	tgagcgtcag	accccgtaga	aaagatcaaa	ggatcttctt	gagatccttt	12480
ttttctgcgc	gtaatctgct	gcttgcaaac	aaaaaaacca	ccgctaaccag	cggtggtttg	12540
tttgccggat	caagagctac	caactctttt	tccgaaggta	actggcttca	gcagagcgca	12600
gataccaaat	actgtccttc	tagtgtagcc	gtagttaggc	caccacttca	agaactctgt	12660
agcaccgcct	acatacctcg	ctctgctaata	cctgttacca	gtggctgctg	ccagtggcga	12720
taagtctgtg	cttaccgggt	tggactcaag	acgatagtta	ccggataagg	cgagcggtc	12780
gggctgaacg	gggggttcgt	gcacacagcc	caacttgag	cgaacgacct	acaccgaact	12840
gagataccta	cagcgtgagc	tatgagaaag	cgccacgctt	cccgaaggga	gaaaggcgga	12900
caggatccg	gtaagcggca	gggtcggaac	aggagagcgc	acgagggagc	ttccaggggg	12960
aaacgcctgg	tatctttata	gtcctgtcgg	gtttcgccac	ctctgacttg	agcgtcgatt	13020
tttgtgatgc	tcgtcagggg	ggcggagcct	atggaaaaac	gccagcaacg	cggccttttt	13080
acggttccctg	gccttttgct	ggccttttgc	tcacatgttc	tttctcggt	tatcccttga	13140
ttctgtggat	aaccgtatta	ccgcctttga	gtgagctgat	accgctcgcc	gcagccgaac	13200
gaccgagcgc	agcaggtcag	tgagcgagga	agcgggaagag	cgccaatac	gcaaaccgcc	13260
tctccccgcg	cgttggccga	ttcattaatg	cagctggcac	gacaggtttc	ccgactggaa	13320
agcgggcagt	gagcgcaacg	caattaatgt	gagttagctc	actcattagg	cacccaggc	13380
tttacacttt	atgcttcggg	ctogtatgtt	gtgtggaatt	gtgagcggat	aacaatttca	13440
cacagaaac	agctatgacc	atgattacgc	caagcgcgca	attaaccctc	actaaaggga	13500
acaaaagctg	tcgagatcta	gatatcgatg	gccatagagt	tacg		13544

<210> 7

<211> 12739

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Rosa 26
targeting vector pOMP1 with TK

<400> 7

ctagggataa	cagggtaata	tagccgcggc	aggccctccg	agcgtggtgg	agccgttctg	60
tgagacagcc	gggtacgagt	cgtgacgctg	gaaggggcaa	gcgggtggtg	ggcaggaaatg	120
cggtccgccc	tgacgaacc	ggagggggag	ggagaaggga	gcggaaaagt	ctccaccgga	180
cgcgcccatg	gctcgggggg	gggggggcag	cggaggagcg	cttccggccg	acgtctcgtc	240
gctgattggc	ttcttttcc	ccgcgcgtgt	gtgaaaacac	aaatggcggtg	ttttggttgg	300
cgtaaggcgc	ctgtcagtta	acggcagccg	gagtgcgcag	ccgcgggcag	cctcgctctg	360
cccactgggt	ggggcgggag	gtaggtgggg	tgaggcgagc	tggacgtgog	ggcgcggtcg	420
gcctctggcg	ggcgggggg	gggagggag	ggtcagcgaa	agtagctcgc	gcgcgagcgg	480
ccgccacccc	ttcccttcc	ctgggggagt	cgttttacc	gccgccggcc	gggectcgtc	540
gtctgattgg	ctctcggggc	ccagaaaact	ggccttgc	attggctcgt	gttcgtgcaa	600
gttgagtcca	tccgcgggcc	agcggggggc	gcgaggaggc	gctcccaggt	tccggccctc	660
ccctcgcccc	cgcgcgcgag	agtctggccg	cgcgccctg	cgcaacgtgg	caggaagcgc	720
gcgctggggg	cggggacggg	cagttaggct	gagcggctgc	ggggcggggtg	caagcacggt	780
tccgacttga	gttgccctcaa	gagggggcgtg	ctgagccaga	cctccatcgc	gcactccggg	840
gagtgagggg	aaggagcgag	ggctcagttg	ggctgttttg	gaggcaggaa	gcacttgctc	900
tccaaaagtc	gctctgagtt	gttatcagta	aggagagctgc	agtggagtag	gcggggagaa	960
ggcgcacccc	ttctccggag	gggggagggg	agtgttgcaa	tacctttctg	ggagttctct	1020
gctgcctcct	ggcttctgag	gaccgccttg	ggcctgggag	aatcccttcc	ccctcttccc	1080
tcgtgatctg	caactccagt	ctttctaggt	aaccgatatac	cctgcagggg	tgacctgcac	1140
gtctagggcg	cagtagtcca	gggtttcctt	gatgatgtca	tacttatcct	gtcccttttt	1200
tttccacagc	tcgcgggtga	ggacaaaact	ttcgcggtct	ttccagtact	cctgcagggtg	1260
actgactgag	tcgactaggg	gtaaccataa	cttcgtataa	tgtatgctat	acgaagtatt	1320
ttataatcta	gaactagtgg	atccacgatt	cgagggcccc	tgcaagtcaa	ttctaccggg	1380
taggggaggc	gcttttccca	aggcagctctg	gagcatgcgc	tttagcagcc	ccgctgggca	1440
cttgcgctca	cacaagtggc	ctctggcctc	gcacacattc	cacatccacc	ggtaggcgcc	1500
aaccggctcc	gttcttttgt	ggcccttctg	cgccaccttc	tactcctccc	ctagtcaggga	1560
agttcccccc	cgcgccgcag	ctcgcgctgt	gcaggacgtg	acaaatggaa	gtagcacgtc	1620
tcaactagtct	cgtgcagatg	gacagcaccg	ctgagcaatg	gaagcgggta	ggcctttggg	1680
gcagcgccca	atagcagctt	tgctccttcg	ctttctgggc	tcagaggctg	ggaagggtg	1740
ggtccggggg	cgggctcagg	gcggcgggcg	ggcgcccgaa	ggcctccggg	ggcctccggg	1800
aggcccgcca	ttctgcacgc	ttcaaaagcg	cacgtctgcc	gcgctgttct	cctcttcttc	1860

atctccgggc	ctttcgacct	gcagccaata	tgggatcggc	cattgaacaa	gatggattgc	1920
acgcaggttc	tccggccgct	tgggtggaga	ggctattcgg	ctatgactgg	gcacaacaga	1980
caatcggtg	ctctgatgcc	gocgtgttcc	ggctgtcagc	gcaggggccc	ccggttcttt	2040
ttgtcaagac	cgacctgtcc	ggtgccctga	atgaactgca	ggacgaggca	gcgcggctat	2100
cgtggctggc	cacgacgggc	gttccttgcg	cagctgtgct	cgacgttgtc	actgaagcgg	2160
gaagggactg	gctgctattg	ggcgaagtgc	cggggcagga	tctcctgtca	tctcaccttg	2220
ctcctgccga	gaaagtatcc	atcatggctg	atgcaatgcg	gcggctgcat	acgcttgatc	2280
cggctacctg	cccatctgac	caccaagcga	aacatcgcat	cgagcgagca	cgtactcgga	2340
tgggaagccg	tcttgctgat	caggatgato	tggacgaaga	gcacagggg	ctcgcgccag	2400
ccgaactgtt	cgccaggctc	aaggcgcgca	tgcgccgacg	cgaggatctc	gtcgtgaccc	2460
atggcgatgc	ctgcttgccg	aatatcatgg	tggaaaatgg	ccgcttttct	ggattcatcg	2520
actgtggcgg	gctgggtgtg	gcggaccgct	atcaggacat	agcgttggct	acccgtgata	2580
ttgctgaaga	gcttggcggc	gaatgggctg	accgcttctc	cgtgctttac	ggtatcgccg	2640
ctcccgaattc	gcagcgcac	gccttctatc	gccttcttga	cgagtctctc	tgaggggata	2700
gatccgctgt	aagtctgcag	aaattgatga	tctattaaac	aataaagatg	tccactaaaa	2760
tgggaagttt	tcctgtcata	ctttgttaag	aagggtgaga	acagagtacc	tacattttga	2820
atggaaggt	tggagctacg	gggtggggg	tgggtggga	ttagataaat	gcctgctctt	2880
tactgaaggc	tctttactat	tgctttatga	taatgtttca	tagttggata	tcataattta	2940
aacaagcaaa	accaaattaa	gggccagctc	attcctccca	ctcatgatct	atagatctat	3000
agatctctcg	tgggatcatt	gtttttctct	tgattccccc	tttgtggttc	taagtactgt	3060
ggttttccaaa	tgtgtcagtt	tcatagcctg	aagaacgaga	tcagcagcct	ctgttccaca	3120
tacacttcat	tctcagttat	gttttgccaa	gttctaattc	catcagaagc	tgactctaga	3180
tcttgcaagg	attcatatga	taacttcgta	taatgtatgc	tatacgaagt	tatggcgcgc	3240
cggtaaccga	agttcctata	ctttctagag	aataggaact	tcggaatagg	aacttcaagc	3300
ttaagcgcta	gaagatgggc	gggagtcttc	tgggcaggct	taaaggctaa	cctgggtgtg	3360
gggcgttgtc	ctgcagggga	attgaacagg	tgtaaaattg	gagggacaag	acttcccaca	3420
gattttcggg	tttgtcggga	agttttttaa	tagggcaaaa	taaggaaaat	gggagggatg	3480
gtagtcatct	ggggttttat	gcagcaaaa	tacaggttat	tattgcttgc	gatccgcctc	3540
ggagtatttt	ccatcgaggt	agattaaaga	catgctcacc	cgagttttat	actctcctgc	3600
ttgagatcct	tactacagta	tgaattaca	gtgtcgcgag	ttagactatg	taagcagaat	3660
tttaatcatt	tttaaagagc	ccagtacttc	atatccattt	ctcccgcctc	ttctgcagcc	3720
ttatcaaaag	gtattttaga	acactcattt	taccccat	ttcatttatt	atactggctt	3780
atccaacccc	tagacagagc	attggcattt	tgcccttctc	gatccttagaa	gtctgatgac	3840
tcatgaaacc	agacagatta	gttacataca	ccacaaatcg	aggctgtagc	tggggcctca	3900
acactgcagt	tcttttataa	ctccttagta	cactttttgt	tgatcctttg	ccttgatcct	3960
taattttcag	tgtctatcac	ctctcccgtc	agtgggtgtc	cacatttggg	cctatttctca	4020
gtccagggag	ttttacaaca	atagatgtat	tgagaatcca	acctaagct	taactttcca	4080
ctcccatgaa	tgctctctc	ctttttctoc	atttataaac	tgagctatta	accattaatg	4140
gttccagggtg	gatgtctcct	ccccatatta	cctgatgtat	cttacatatt	gccaggctga	4200
tatttttaaga	cattaaaagg	tatatttcat	tattgagcca	catggatttg	attactgctt	4260
actaaaattt	tgcatgtgta	cacatctgta	aaagggtggt	ccttttggaa	tgcaaagttc	4320
aggtgtttgt	tgtctttcct	gacctaaggt	cttgtgagct	tgtatttttt	ctattttaagc	4380
agtgttttct	cttggactgg	cttgactcat	ggcatttctac	acgttattgc	tggcttaaat	4440
gtgattttgc	caagcttctc	caggacctat	aattttgtct	gacttgtago	caaacacaa	4500
taaaatgatt	aagcaacaaa	tgtattttgt	aagcttgggt	tttaggttgt	tgtgttgtgt	4560
gtgcttgtgc	tctataataa	tactatccag	gggctggaga	ggtggctcgg	agttcaagag	4620
cacagactgc	tcttccagaa	gtcctgagtt	caattcccag	caaccacatg	gtggctcaca	4680
accatctgta	atgggatctg	atgcctctct	ctgggtgtgc	tgaagaccac	aagtgtattc	4740
acattaaata	aataaatcct	ccttcttctt	cttttttttt	tttttaaaga	gaataactgc	4800
tccagtagaa	tttactgaag	taatgaaata	ctttgtgttt	gttccaatat	ggtagccaat	4860
aatcaaatta	ctctttaagc	actggaaatg	ttaccaagga	actaattttt	atttgaagtg	4920
taactgtgga	cagaggagcc	ataactgcag	acttgtggga	tacagaagac	caatgcagac	4980
tttaattgtct	tttctcttac	actaagcaat	aaagaaataa	aaattgaact	tctagtatcc	5040
tatttgttta	aactgctagc	tttacttaac	ttttgtgctt	catctataca	aagctgaaa	5100
ctaagtctgc	agccattact	aaacatgaaa	gcaagtaatg	ataatttttg	atttcaaaaa	5160
tgtagggcca	gagtttagcc	agccagtggt	ggtgcttgcc	tttatgcctt	taatcccagc	5220
actctggagg	cagagacagg	cagatctctg	agtttgagcc	cagcctggct	tacacatcaa	5280
gttctatcta	ggatagccag	gaatacacac	agaaaccctg	ttggggaggg	gggctctgag	5340
atttcataaa	attataattg	aagcattccc	taattgagcca	ctatggatgt	ggctaaatcc	5400
gtctaccttt	ctgatgagat	ttgggtatta	tttttctgt	ctctgctgtt	gggtgggtct	5460
tttgacactg	tgggctttct	ttaaagcctc	cttcctgcca	tgtggtctct	tggttgctac	5520
taacttccca	tggcttaaat	ggcatggctt	tttgcttctt	aagggcagct	gctgagattt	5580
gcagcctgat	ttccagggtg	gggttgggaa	atctttcaaa	cactaaaatt	gtcctttaat	5640
ttttttttta	aaaaatgggt	tatataataa	acctcataaa	atagttatga	ggagtggagt	5700
ggactaatat	taaatgagtc	cctcccctat	aaaagagcta	ttaaggcttt	ttgtcttata	5760
cttaactttt	tttttaaatg	tggtatcttt	agaaccaagg	gtcttagagt	tttagtatat	5820
agaaactggt	gcacgcctta	atcagatttt	ctagtttcaa	atccagagaa	tccaaattct	5880

tcacagccaa	agtcaaatta	agaattttctg	acttttaaatg	ttaattttgct	tactgtgaat	5940
ataaaaaatga	tagctttttcc	tgaggcaggg	tctcactatg	tatctctgcc	tgatctgcaa	6000
caagatatgt	agactaaagt	tctgcctgct	tttgtctcct	gaatactaag	gttaaaatgt	6060
agtaatactt	ttggaacttg	caggtcagat	tcttttatag	gggacacact	aaggggagctt	6120
gggtgatagt	tggtaaaatg	tgtttcaagt	gatgaaaact	tgaattatta	tcaccgcaac	6180
ctactttttta	aaaaaaaaaag	ccaggcctgt	tagagcatgc	ttaagggatc	cctaggactt	6240
gctgagcaca	caagagtagt	tacttggcag	gctcctgggtg	agagcatatt	tcaaaaaaca	6300
aggcagacaa	ccaagaaact	acagttaagg	ttacctgtct	ttaaaccatc	tgcatataca	6360
cagggatatt	aaaatattcc	aaataatatt	tcattcaagt	tttcccccot	caaattggga	6420
catggatttc	tccggtgaat	aggcagagtt	ggaaactaaa	caaatgttgg	ttttgtgatt	6480
tgtgaaattg	ttttcaagtg	atagttaaag	cccatgagat	acagaacaaa	gctgctattt	6540
cgagggtctct	ttggtttatac	tcagaagcac	ttctttgggt	ttccctgcac	tatcctgac	6600
atgtgctagg	cctaccttag	gctgattggt	gttcaaataa	acttaagttt	cctgtcaggt	6660
gatgtcatat	gatttcatat	atcaaggcaa	aacatgttat	atatgttaaa	catttgtact	6720
taatgtgaaa	gttaggtctt	tgtgggtttg	atttttaatt	ttcaaaacct	gagctaaata	6780
agtcattttt	acatgtctta	catttgggtg	aattgtataa	ttgtggtttg	caggcaagac	6840
tctctgacct	agtaacccta	cctatagagc	actttgctgg	gtcacaagtc	taggagtcaa	6900
gcatttcacc	ttgaagttga	gacgttttgt	tagtgtatac	tagtttatat	gttggaggac	6960
atgtttatcc	agaagatatt	caggactatt	tttgactggg	ctaaggaatt	gattctgatt	7020
agcaactgta	gtgagcattg	agtggccttt	aggcctgaat	tggagtcaat	tgtatatctc	7080
aaataatgct	ggcctttttt	aaaaagcctt	tgctctttat	cacctgtttt	tctacataat	7140
ttttgttcaa	agaaataactt	gtttggatct	ccttttgaca	acaatagcat	gttttcaagc	7200
catatttttt	ttcctttttt	tttttttttt	tggtttttcg	agacaggggt	tctctgtata	7260
gccctggctg	tcttggaact	cactttgtag	accaggctgg	cctcgaactc	agaaatccgc	7320
ctgcctctgc	ctcctgagtg	ccgggattaa	aggcgtgcac	caccacgcct	ggctaagttg	7380
gatattttgt	tatataacta	taaccaatac	taactccact	gggtggattt	ttaattcagt	7440
cagtagtctt	aagtggctct	tattggcctt	tcattaaaa	ctactgttca	ctctaacaga	7500
ggctgttgg	actagtggca	cttaagcaac	ttcctacgga	tatactagca	gattaagggt	7560
cagggataga	aactagtcta	gcgttttgta	tacctaaccag	ctttatacta	ccttgttctg	7620
atagaaatat	ttcaggacat	ctagcacgtg	ttaactcgag	ctgcaggatt	cgagggcccc	7680
ggcaggtcaa	ttctaaccggg	taggggaggg	gcttttccca	aggcagctcg	gagcatcgcc	7740
tttagcagcc	ccgctgggca	cttgccgcta	cacaagtggc	ctctggcctc	gcacacattc	7800
cacatccacc	ggtaggcgcc	aaccggctcc	gttctttgg	ggccctctcg	cgccaccttc	7860
tactcctccc	ctagtccagg	agttccccc	cgccccgcag	ctcgcgtcgt	gcaggacgtg	7920
acaaatggaa	gtagcacgtc	tactagtct	cgtgcagatg	gacagcaccg	ctgagcaatg	7980
gaagcgggta	ggcctttggg	gcagcggcca	atagcagctt	tgctccttcg	ctttctgggc	8040
tcagaggctg	ggaaggggtg	ggtccggggg	cgggtccagg	ggcgggctca	ggggcggggc	8100
gggcgccga	aggtcctccg	gaggcccgcc	attctgcacg	cttcaaaagc	gcacgtctgc	8160
cgcgctgttc	tcctcttctt	catctccggg	cctttcgacc	tgacagcaat	gcaccgtcct	8220
tgccatcatg	gcctcgtacc	ccggccatca	acacgcgtct	gcgttcgacc	aggctgcgcg	8280
ttctcgcgcc	catagcaacc	gacgtacggc	gttgccgctt	cgccggcagc	agaagccac	8340
ggaagtccgc	ccggagcaga	aaatgcccac	gctactgcgg	gtttatatag	acggtcccca	8400
cgggatgggg	aaaaccacca	ccacgcaact	gctggtggcc	ctgggttcgc	gcgacgatat	8460
cgtctacgta	cccagagccga	tgacttactg	gcgggtgctg	ggggcttccg	agacaatcgc	8520
gaacatctac	accacacaac	accgcctcga	ccagggtgag	atatcgcccg	gggacgcggc	8580
ggtggttaatg	acaagcggcc	agataacaat	gggcatgcct	tatgccgtga	ccgacgcctg	8640
tctggctcct	catatcgggg	gggaggtcgt	gagctcacat	gccccgcccc	cgccctcac	8700
cctcatcttc	gaccgcatc	ccatcgccgc	cctcctgtgc	taccggcgcc	cgcggtacct	8760
tatgggcagc	atgaccccc	aggccgtgct	ggcgttcgtg	gcccctcatcc	cgccgacctt	8820
gcccggcacc	aacatcgtgc	ttggggccct	tccggaggac	agacacatcg	accgcctggc	8880
caaacgccag	cgccccggcg	agcggctgga	cctggctatg	ctggctgcga	ttcgccgcgt	8940
ttacgggcta	cttgccaata	cggtgcggtg	tctgcagtg	ggcgggtcgt	ggcgggagga	9000
ctggggacag	ctttcgggga	cggccgtggc	gccccagggt	gcccagcccc	agagcaacgc	9060
gggcccacga	ccccatatcg	gggacacgtt	atttaccctg	tttcgggccc	ccgagttgct	9120
ggcccccaac	ggcgacctgt	ataacgtgtt	tgctggggcc	ttggacgtct	tgcccaaacg	9180
cctccgttcc	atgcacgtct	ttatcctgga	ttacgaccaa	tcgccccgcg	gctgccggga	9240
cgccctgctg	caacttacct	ccgggatggt	ccagaccac	gtcaccaccc	ccggctccat	9300
accgacgata	tgcgacctgg	cgcgcacgtt	tgccccggag	atggggagg	ctaactgagg	9360
ggatcgatcc	gtcctgtaag	tctgcagaaa	ttgatgatct	attaaacaat	aaagatgtcc	9420
actaaaaatg	aagtttttcc	tgtcatactt	tgtaaagaag	ggtgagaaca	gagtacctac	9480
attttgaatg	gaaggattgg	agctacgggg	gtgggggtgg	ggtgggatta	gataaatgcc	9540
tgctctttac	tgaaggctct	ttactattgc	tttatgataa	tgtttcatag	ttggatatca	9600
taattttaaac	aagcaaaaacc	aaattaaagg	ccagctcatt	cctccactc	atgatctata	9660
gatctataga	tctctcgtgg	gatcattggt	tttctcttga	ttcccacttt	gtggttctaa	9720
gtactgtggt	ttccaaatgt	gtcagtttca	tagcctgaag	aacgagatca	gcagcctctg	9780
ttccacatac	acttcattct	cagtattggt	ttgccaagtt	ctaattccat	cagaagctga	9840
ctctaggccg	agctccaatt	cgccctatag	tgagtcgtat	tacaattcac	tggccgctcg	9900

```

tttacaacgt cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc ttgcagcaca 9960
tccccctttc gccagctggc gtaatagcga agaggccgc accgatcgcc cttcccaaca 10020
gttgcgcagc ctgaatggcg aatgggacgc gccctgtagc ggcgcattaa gcgcggcggg 10080
tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt 10140
cgctttcttc ccttcctttc tcgccacgtt cgccggttt ccccgtaag ctctaaatcg 10200
ggggctccct ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga 10260
ttaggggtgat ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac 10320
gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc 10380
tatctcggtc tattcttttg atttataagg gattttgccg atttcggcct attggttaaa 10440
aaatgagctg atttaacaaa aatttaacgc gaattttaac aaaaatttaa cgcttacaat 10500
ttagggtggc tttttcgggg aaatgtgcgc ggaacccta tttgtttatt tttctaaata 10560
cattcaataa tgtatccgct catgagacaa taacctgat aaatgcttca ataatttga 10620
aaaaggaaga gtatgagtat tcaacatttc cgtgtcgcgc ttattccctt ttttgccgca 10680
ttttgccttc ctgtttttgc tcaccagaa acgctggtga aagtaaaaga tgctgaagat 10740
cagttgggtg cagcagtggtg ttacatcgaa ctggatctca acagcggtaa gatccttgag 10800
agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagtctt gctatgtggc 10860
gcggtattat ccgctattga ccgcgggcaa gagcaactcg gtccgcgcac acactattct 10920
cagaatgact tgggtgagta ctaccagtc acagaaaagc atcttacgga tggcatgaca 10980
gtaagagaat tatgcagtgc tgccataacc atgagtata acactgcggc caacttactt 11040
ctgacaacga tcggaggacc gaaggagcta accgctttt tgcacaacat gggggatcat 11100
gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt 11160
gacaccacga tgcctgtagc aatggcaaca acgttgcgca aactattaac tggcgaacta 11220
cttactctag cttcccgcca acaattaata gactggatgg aggcggataa agttgcagga 11280
ccacttctgc gctcggccct tcgggtggc tgggtttatt ctgataaatc tggagccggt 11340
gagcgtgggt ctgcgggtat cattgcagca ctggggccag atggtaagcc ctcccgatc 11400
gtagttatct acacgacggg gactcaggca actatggatg aacgaaatag acagatcgct 11460
gagataggtg cctcactgat taagcattgg taactgtcag accaagttta ctcatatata 11520
ctttagattg atttaaaact tcatttttaa tttaaaagga tctagtgtaa gatccttttt 11580
gataatctca tgacaaaat cccttaacgt gatttttcgt tccactgagc gtcagacccc 11640
gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgccgctaact ctgctgcttg 11700
caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact 11760
ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt ccttctagt 11820
tagccgtagt tagccaacca cttcaagaac cctgtagcac cgcctacata cctcgctctg 11880
ctaactcgtg taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttgga 11940
tcaagacgat agttaccgga taaggcgag cggtcgggct gaacggggg ttcgtgcaca 12000
cagcccagct tggagcgaa cgcctacacc gaactgagat acctacagcg tgagctatga 12060
gaaagcgcca cgcttcccga agggagaaa ggcgacaggt atccggtaag cggcagggc 12120
ggaacaggag agcgacgag ggagcttcca ggggaaaacg cctggtatct ttatagctc 12180
gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcg 12240
agcctatgga aaaacgccag caacgcggcc tttttacggt tccgtgacct ttgctggcct 12300
tttgctcaca tgttctttcc tgcggtatcc cctgattctg tggataaccg tattaccgcc 12360
tttgagttag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagttagc 12420
gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg gccgattcat 12480
taatgcagct ggcacgacag gtttcccgcac tggaaagcgg gcagtgcagc caacgcaatt 12540
aatgtgagtt agctcactca ttaggcaccc caggctttac actttatgct tccggctcgt 12600
atgtgtgtg gaattgtgag cggataacaa tttcacacag gaacacagta tgaccatgat 12660
tacgccaagc gcgcaattaa ccctcactaa agggaacaaa agctgtcgag atctagatat 12720
cgatggccat agagttacg 12739

```

<210> 8

<211> 15342

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Rosa 26 locus targeted with targeting vector

<400> 8

```

aagcttctca cgtagcaacc agagctccag agccagcagc tgctgccgdc ttgtatactc 60
actcctgtga tccaacacag gagcaacctt ttctttaccc caccgccact tcttaacaca 120
cttttttttg gggggggggg gggaacaaat gctccatgct ggaaggattg gaactatgct 180
tttagaaaag aacaatccta aggtcacttt taaattgagg tctttgattt gaaaatcaac 240
aaataccaaa ttccaaatat tcgttttaat taaaccagca atgtggatat aagcattaag 300
ttttagtttt aaaaaggcca attttccaaa cattcagcaa tcatatttaa atttacagct 360
aggaacaaga gccttgggtc atgtcctacc aaagaacata actcaatatt ctacacatga 420

```

caatctgaat	aaccttaaaag	cctctaatacc	cataacaggc	cacaaattttt	ggacagagaa	480
ctaatagatcc	tccttgagaaa	actggaagaa	atccagggaa	aagaaattcc	tgtgtcctcc	540
aaactcagaa	atctctaatt	atgtcagtat	tctctgcttt	agtcctaggt	cagattgcac	600
acatctaaaa	taacctctta	aagttttcct	cctagcgacc	taaaccatta	ttaatatcaa	660
attaaccatc	aaaaactttt	cctctcaata	tgctgcacac	aaacctctc	ctggaacctc	720
ctccatctgg	atcctcccca	atcaaaagta	taggtattta	acataataagc	aaggaagtaa	780
tgtaaacatg	accttggtca	caaataatgtc	atctaaaaac	aatttagtca	aggtagggag	840
gaaattcgag	aacctgaatc	tttttaagta	ttttgagcac	aggaacaatt	ggcaaaagga	900
atccaggtat	agacaaaacc	cagagcccag	agctctgggc	gaaaaatgag	ttgctggtga	960
agacgttaca	caagtaacat	gagaaagcag	aaaatgcagg	tcatccacgc	acccctgacc	1020
caggccagca	ggcggggctg	cagcatcagt	acacaggaga	aagatcctta	ttcctaagaa	1080
tgagaaaggc	aaagcgccc	gatagaataa	attagcatag	aaggggcttt	cccaggagt	1140
aaaacttttc	ttctgagcga	ttacctacta	aaaccagggc	ttttgccac	taccatttac	1200
ctaggatctt	ggcttgacag	gattcatagg	ggcatatccc	tccccctctt	ctttagagtc	1260
gttcttaaaa	gatcgctctc	cacgccttag	gcagggaaaa	cgacaaaatc	tggtcaatt	1320
ccaggctaga	acctacaaa	ttcaacaggg	atatcgcaag	gatactgggg	catagccac	1380
agggagtcca	agaatgtgag	gtgggggtg	cgaaggtaat	gtctttggtg	tgggaaaagc	1440
agcagccatc	tgagatagga	actggaaaac	cagaggagag	gcgttcagga	agattatgga	1500
ggggaggact	gggccccac	gagcgaccag	agttgtcaca	aggccgcaag	aacaggggag	1560
gtggggggct	cagggacaga	aaaaaaaaagta	tgtgtatttt	gagagcaggg	ttgggaggcc	1620
tctcctgaaa	agggataaaa	cgtggagtag	gcaataccca	ggcaaaaagg	ggagaccaga	1680
gtagggggag	gggaagagtc	ctgaccagc	gaagacatta	aaaaggtagt	ggggtcgact	1740
agatgaaggga	gagcctttct	ctctgggcaa	gagcgggtca	atgggtgtga	aaggtagctg	1800
agaagacgaa	aagggcaagc	atcttctctgc	taccaggctg	gggaggccca	ggcccacgac	1860
cccaggagga	gggaacgcag	ggagactgag	gtgacccttc	tttcccccg	ggcccggctg	1920
tgtggttcgg	tgtctctttt	ctgttgacc	cttaccttga	cccaggcgct	gcccgggctt	1980
gggcccgggc	tgcggcgac	ggcactccc	ggaggcagcg	agactcgagt	taggcccac	2040
gcggcgccac	ggcgtttcct	ggcgggaat	ggcccgtacc	cgtgaggtgg	gggtgggggg	2100
cagaaaaggc	ggagcgagcc	cgagcgggga	gggggagggc	caggggaggga	gggggagggc	2160
actactgtgt	tggcggactg	gcgggactag	ggctgcgtga	gtctctgagc	gcaggcgggc	2220
ggcgccggcc	cctcccccg	cgccggcagc	ggcgccagcg	gcggcagctc	actcagccc	2280
ctgcccagc	ggaaacgcca	ctgaccgcac	ggggattccc	agtgcggcg	ccaggggcac	2340
gcgggacacg	cccctccc	cgcgcctt	ggcctctccg	cccacggccc	cacacttatt	2400
ggccggtgcg	ccgccaatca	gcggaggctg	ccggggccgc	ctaaagaaga	ggctgtgctt	2460
tggggctccg	gctcctcaga	gagcctcgcc	taggtagggg	atcgggactc	tggcgaggag	2520
gcggcttggt	gcgtttgcgg	ggatggcg	ccgcggcagg	ccctccgagc	gtggtggagc	2580
cgttctgtga	gacagccggg	tacgagtcgt	gacgctggaa	ggggcaagcg	ggtggtgggc	2640
aggaatgcgg	tccgcctgc	agcaaccgga	gggggaggga	gaagggagcg	gaaaagctc	2700
caccggagcg	cgccatggct	cggggggggg	ggggcagcg	aggagcgctt	ccggccgacg	2760
tctcgtcgct	gattggcttc	ttttctctcc	gccgtgtgtg	aaaacacaaa	tggcgtgttt	2820
tggttggcgt	aaggcgctg	tcagttaacg	gcagccggag	tgccgagccg	ccggcagcct	2880
cgctctgccc	actgggtggg	gcgggaggta	ggtgggggtga	ggcgagctgg	acgtggggg	2940
gcggtcggcc	tctggcgggg	cgggggaggg	gagggagggt	cagcgaaaagt	agctgcgcg	3000
cgngcgccg	cccaccctcc	ccttctctg	ggggagtcgt	tttaccgcgc	gcccggcggg	3060
cctcgtcgct	tgattggctc	tcggggccca	gaaaactggc	ccttgccatt	ggctcgtgtt	3120
cgtgcaagtt	gagtcacatc	gccggccagc	gggggaggcg	aggaggcgct	cccaggttcc	3180
ggccctcccc	tcggccccgc	gccgcagagt	ctggccgcgc	gcccctgcgc	aacgtggcag	3240
gaagcgcgcg	ctggggggcg	ggacgggcag	tagggctgag	cggctgcggg	gcgggtgcaa	3300
gcacgtttcc	gacttgagtt	gcctcaagag	ggcggtgctg	agccagacct	ccatcgcgca	3360
ctccggggag	tggagggaag	gagcgagggc	tcagttgggc	tgttttggag	gcaggaagca	3420
cttgcctctc	caaagtcgct	ctgagttggt	atcagtaagg	gagctgcagt	ggagtaggcg	3480
gggagaaggc	cgccaccttc	tcgggagggg	ggaggggagt	gttgcaatac	ctttctggga	3540
gttctctgct	gcctcctggc	ttctgaggac	cgccctgggc	ctgggagaat	cccttcccc	3600
tcttcctctg	tgatctgcaa	ctccagttct	tctaggtaac	cgatatccct	gcagggggtga	3660
cctgcacgtc	tagggcgag	tagtccaggg	tttctctgat	gatgtcatac	ttatcctgtc	3720
cctttttttt	ccacagctcg	cgggttagga	caaactcttc	gcggctcttc	cagtactcct	3780
gcaggtgact	gactgagtcg	actaggggta	accataactt	cgtataatgt	atgctatacg	3840
aagttattta	taatctagaa	ctagtggtatc	cacgattcga	gggcccctgc	aggtcaattc	3900
taccgggtag	gggaggcgct	tttcccaagg	cagtcctggag	catgcgcttt	agcagcccc	3960
ctgggcactt	ggcgctacac	aagtggcctc	tggcctcgca	cacattccac	atccaccggt	4020
aggcgccaac	cggctccgtt	ctttgggtggc	cccttcgcgc	cacottctac	tcctccccta	4080
gtcaggaagt	tcctcccgcc	cccgagctc	gcgtcgtgca	ggacgtgaca	aatggaagta	4140
gcacgtctca	ctagtctcgt	gcagatggac	agcaccgctg	agcaatggaa	gcgggtaggc	4200
ctttggggca	gcggccaata	gcagctttgc	tccttcgctt	tctgggctca	gaggtgggga	4260
aggggtgggt	ccggggggcg	gctcagggcg	ggctcagggg	cggggcgggc	gcccgagggt	4320
cctccggagg	ccggcattc	tgcacgcttc	aaaagcgac	gtctgcgcg	ctgttctcct	4380
cttctctatc	tcggggcctt	tcgacctgca	gccaatatgg	gatcgcccat	tgaacaagat	4440

ggattgcacg	cagggttctcc	ggcgcgttgg	gtggagaggc	tattcggtta	tgactgggca	4500
caacagacaa	tcggctgctc	tgatgccgcc	gtgttcgggc	tgtagcgca	ggggcgcccg	4560
gttctttttg	tcaagaccga	cctgtccggg	gccctgaatg	aactgcagga	cgaggcagcg	4620
cggctatcgt	ggctggccac	gacgggcgtt	ccttgccgag	ctgtgctcga	cggtgtcact	4680
gaagcgggaa	gggactggct	gctattgggc	gaagtgcggg	ggcaggatct	cctgtcatct	4740
caccttgctc	ctgccgagaa	agtatccatc	atggctgatg	caatgcggcg	gctgcatacg	4800
cttgatccgg	ctacctgccc	attcgaccac	caagcgaaac	atcgcatcga	gcgagcacgt	4860
actcggatgg	aagccgggtc	tgtcgatcag	gatgatctgg	acgaagagca	tcaggggctc	4920
gcgccagccg	aactgttcgc	caggctcaag	gcgcgcagtc	ccgacggcga	ggatctcgtc	4980
gtgacccatg	gogatgcctg	cttgccgaat	atcatgggtg	aaaatggccg	cttttctgga	5040
ttcatcgact	gtggccggct	gggtgtggcg	gacogctatc	aggacatagc	gttggctacc	5100
cgtgatattg	ctgaagagct	tgccggcgaa	tgggctgacc	gcttcctcgt	gctttacggg	5160
atcgccgctc	ccgattcgca	gcgcacgcgc	ttctatcgcc	ttcttgacga	gttcttctga	5220
ggggatcgat	ccgctgtaag	tctgcagaaa	ttgatgatct	attaaacaat	aaagatgtcc	5280
actaaaatgg	aagtttttcc	tgtcatactt	tgtaagaag	ggtgagaaca	gagtacctac	5340
attttgaatg	gaaggattgg	agctacgggg	gtgggggtgg	ggtgggatta	gataaatgcc	5400
tgctctttac	tgaaaggctc	ttactattgc	tttatgataa	tgtttcatag	ttggatatca	5460
taatttaaac	aagcaaaacc	aaattaaggg	ccagctcatt	cctcccactc	atgatctata	5520
gatctataga	tctctcgtgg	gatcattgtt	tttctcttga	ttcccacttt	gtggttctaa	5580
gtactgtggt	ttccaaatgt	gtcagtttca	tagcctgaag	aacgagatca	gcagcctctg	5640
ttccacatac	acttcattct	cagtattgtt	ttgccaaagt	ctaattccat	cagaagctga	5700
ctctagatcc	tgagggaatt	catatgataa	cttcgtataa	tgtagctat	acgaagtatt	5760
ggcgcgcggg	taaccgaagt	tcctatactt	tctagagaat	aggaacttcg	gaataggaac	5820
ttcaagctta	agcgcctagaa	gatgggcggg	agtcttctgg	gcaggcttaa	aggctaacct	5880
ggtgtgtggg	cgttgtcctg	caggggaatt	gaacagggtg	aaaattggag	ggacaagact	5940
teccacagat	tttcggtttt	gtcgggaagt	tttttaatag	gggcaataa	ggaaaaatgg	6000
aggataggtg	gtcatctggg	gttttatgca	gcaaaaactac	aggttattat	tgcttgggat	6060
ccgcctcgga	gtattttcca	tcgaggtaga	ttaaagacat	gctcaccgga	gttttatact	6120
ctcctgcttg	agatccttac	tacagtatga	aattacagtg	tcgcgagtta	gactatgtaa	6180
gcagaatttt	aatcattttt	aaagagccca	gtacttcata	tccattttct	ccgctccttc	6240
tgagccttta	tcaaaaggta	ttttagaaca	ctcatttttag	ccccattttc	atltattata	6300
ctggcttatc	caacccttag	acagagcatt	ggcatttttc	ctttcctgat	cttagaagtc	6360
tgatgactca	tgaaaccaga	cagattagtt	acatacacca	caaatcgagg	ctgtagctgg	6420
ggcctcaaca	ctgcagttct	tttataactc	cttagtacac	ttttgttga	tcctttgcct	6480
tgatccttaa	ttttcagtgt	ctatcacctc	tcocgctcagt	ggtgttccac	atltgggcct	6540
attctcagtc	cagggagttt	tacaacaata	gatgtattga	gaatccaacc	taaagcttaa	6600
ctttccactc	ccatgaatgc	ctctctcctt	tattctocatt	tataaactga	gctattaac	6660
attaatggtt	ccaggtggat	gtctcctccc	catattacct	gatgtatctt	acatattgcc	6720
aggctgatat	tttaagacat	taaaagggtat	atltcattat	tgagccacat	ggtattgatt	6780
actgcttact	aaaattttgt	cattgtacac	atctgtaaaa	ggtggttcc	tttggaatgc	6840
aaagtccagg	tgtttggtgt	ctttcctgac	ctaaggctct	gtgagcttgt	atlttttcta	6900
tttaagcagt	gctttctctt	ggactggctt	gactcatggc	attctacacg	ttattgctgc	6960
tcataatgtg	atlttgccaa	gcttcttcag	gacctataat	tttgcttgac	ttgtagccaa	7020
acacaagtaa	aatgattaag	caacaaatgt	atltgtgaag	cttggttttt	aggttggtgt	7080
gttgtgtgtg	cttgtgctct	ataataatac	tatccagggg	ctggagaggt	ggctcggagt	7140
tcaagagcac	agactgctct	tcagaagtc	ctgagttcaa	ttccagcaa	ccacatggtg	7200
gctcacaacc	atctgtaatg	ggatctgatg	cctctctctg	gtgtgtctga	agaccacaag	7260
tgtattcaca	ttaataaat	aaatcctcct	tttctctctt	tttttttttt	ttaaagagaa	7320
tactgtctcc	actagaattt	actgaagtaa	tgaaatactt	tgtgtttgtt	ccaatatggt	7380
agccaataat	caaattactc	tttaagcact	ggaaatgtta	ccaaggaact	aatttttatt	7440
tgaagtgtaa	ctgtggacag	aggagccata	actgcagact	tgtgggatac	agaagaccac	7500
tgcagacttt	aatgtctttt	ctcttacact	aagcaataaa	gaaataaaaa	ttgaacttct	7560
agtatcctat	ttgtttaaac	tgctagcttt	acttaacttt	tgtgcttcat	ctatacaaa	7620
ctgaaagcta	agtctgcagc	cattactaaa	catgaaagca	agtaatgata	atlttggtat	7680
tcaaaaatgt	agggccagag	tttagccagc	cagtgtggtg	gcttgccctt	atgcctttta	7740
tcccagcact	ctggaggcag	agacaggcag	atctctgagt	ttgagcccag	cctgggtctac	7800
acatcaagtt	ctatctagga	tagccaggaa	tacacacaga	aacctgtgtg	gggagggggg	7860
ctctgagatt	tcataaaaatt	ataattgaag	cattccctaa	tgagccacta	tgatgtggc	7920
taaatccgct	tacctttctg	atgagatttt	ggtattattt	tttctgtctc	tgctgttggg	7980
tggttctttt	gacactgtgg	gctttcttta	aagcctcctt	cctgccatgt	ggtctcttgt	8040
ttgctactaa	cttcccatgg	cttaaatggc	atggcttttt	gccttctaag	ggcagctgct	8100
gagatttgca	gcctgatttc	caggggtggg	ttgggaaatc	tttcaaacac	taaaattgtc	8160
ctttaatttt	ttttttaaaa	aatgggttat	ataataaacc	tcataaaaata	gttatgagga	8220
gtgaggtgga	ctaattattaa	atgagtcctt	cccctataaa	agagctatta	aggctttttg	8280
tcttatagct	aacttttttt	ttaaatgtgg	ctcttttaga	accaaggggtc	ttagagtttt	8340
agtatacaga	aactgttgca	tcgcttaatc	agattttcta	gtttcaaatc	cagagaatcc	8400
aaattcttca	cagccaaagt	caaattaaga	atltctgact	tttaatgtta	atltgcttac	8460

tgtgaatata aaaatgatag cttttcctga ggcagggtct cactatgtat ctctgcctga 8520
tctgcaacaa gatattgtaga cttaaagtctt gcctgctttt gtctcctgaa tactaagggtt 8580
aaaatgtagt aatacttttg gaacttgacag gtcagattctt tttatagggg acacactaag 8640
ggagcttgagg tgatagttgg taaaatgtgt ttcaagtgat gaaaacttga attattatca 8700
ccgcaacctt ctttttaaaa aaaaaagcca ggcctgttag agcatgctta agggatccct 8760
aggacttgct gagcacacaa gagtagttac ttggcaggct cctgggtgaga gcatatttca 8820
aaaaacaagg cagacaacca agaaactaca gtttaagggtta cctgtcttta aaccatctgc 8880
atatacacag ggatattaaa atattccaaa taatatttca ttcaagtttt ccccatcaa 8940
attgggacat ggatttctcc ggtgaatagg cagagttgga aactaaacaa atgttgggtt 9000
tgtgatttgg gaaattgttt tcaagtata gtttaagccc atgagataca gaacaaagct 9060
gctatttcca ggtctcttgg tttatactca gaagcacttc tttgggtttc cctgactat 9120
cctgatcatg tgctaggcct accttaggct gattgttgtt caaataaact taagtttctt 9180
gtcagggtgat gtcatatgat ttcatatatt aaggcaaaac atgttatata ttttaattttc 9240
ttgtacttaa tgtgaaagtt aggtctttgt ggggttgatt ttttaattttc aaaacctgag 9300
ctaaataagt cattttttaca tgtctttacat ttggtggaat tgtataattg tggtttgag 9360
gcaagactct ctgacctagt aacctacct atagagcact ttgctgggtc acaagtctag 9420
gagtcaagca tttcaccttg aagttgagac attttgttag tgtatactag tttatatgtt 9480
ggaggacatg tttatccaga agatattcag gactattttt gactgggcta aggaattgat 9540
tctgattagc actgttagtg agcattgagt ggcttttagg ctggaattgg agtcacttgt 9600
atatctcaaa taatgctggc cttttttaaa aagcccttgt tctttatcac cctgttttct 9660
acataatttt tgttcaaaga aatacttgtt tggatctcct ttgacaaca atagcatgtt 9720
ttcaagccat attttttttt cttttttttt ttttttttgg tttttcgaga cagggtttct 9780
ctgtatagcc ctggctgtcc tggaaactcac tttgtagacc aggtggcct cgaactcaga 9840
aatccgctg cctctgcctc ctgagtgccg ggattaaagg cgtgcaccac cacgcctggc 9900
taagttggat attttgttat ttgtctttat tggcccttca ttaaaacta ctccactggg 9960
attcagtcag tagtcttaag agtggcactt aagcaacttc ctacggatat actagcat 10020
taacagaggc tgttggtact ggttagaaac tagtctagcg tttgtatac ctaccagctt 10080
taagggtcag ggatgaaac tagtctagcg tttgtatac ctaccagctt tatactacct 10140
tgttctgata gaaatatttc aggcactcta gactgtacta taaggttgat ggtaagctta 10200
taaggaactt gaaagtggag taactactcc atttctctga ggggagaatt aaaattttt 10260
accaagtgtt gttgagccac tgagaatggc ctcagaacat aacttcttaa ggaaccttc 10320
cagattgccc tcaacactgc accacatttg gtctctgttg tttgttttt cctcctttag 10380
agtcttaatt aagaatatta attgtgtaatt tttgtttttt tttgtgtctg tctgagcagc 10440
gaggacagga cagtgttgt ttaaggctat atttctgtctg tctgagcagc aacaggtctt 10500
cgagatcaac atgatgttca taatcccaag atgttgccat ttatgttctc agaagcaagc 10560
agaggcatga tggtcagtga cagtaatgtc actgtgttaa atgttgctat gcagtttga 10620
tttttctaatt gtagtgtagg tagaacatat gtgttctgta tgaattaaac tcttaagtta 10680
caccttgat atctcatgca atgtgttatg caattaccat tttaagtatt gtagctttct 10740
ttgtatgtga ggataaagggt gttgtcata aaatgttttg aacatttccc caaagttcca 10800
aattataaaa ccacaacgtt agaacttatt tatgaacaat gggtgtagtt tcatgctttt 10860
aaaatgctta attattcaat taacaccgtt tgtgttataa tatataataa actgacatgt 10920
agaagtgttt gtccagaaca tttcttaaat gtatactgtc tttagagagt ttaatatagc 10980
atgtcttttg caacatacta acttttgtgt tgggtcgagc aatatttgtt agtcattttg 11040
aaaggtatga tttcaatgag tgtcagattg ttttgaatgt tattgaacat tttaaatgca 11100
gacttggtcg tgttttagaa agcaaaactg tcagaagctt tgaactagaa attaaaaagc 11160
tgaagtattt cagaaggga ataaagctact tgctgtatta gttgaaggaa agtgtaatag 11220
cttagaaaaat ttaaaaccat atagtgtgca ttgctgaata tctggcagat gaaaagaaat 11280
actcagtggt tcttttgagc aatataacag cttgttatat taaaaatttt cccacagat 11340
ataaactcta atctataact cataaatgtt acaaatggat gaagcttaca aatgtggctt 11400
gacttgctac tgtgcttgtt ttagttatgt gaaagtgttg caataaacct atgtcctaaa 11460
tagtcaaaact gtggaatgac tttttaatct attggtttgt ctagaacagt tatgttgcca 11520
tttgccctaa tggtgaaaga aaaagtgggg agtgccttg cactgttcat ttgtgggtgt 11580
aaccaaagag gggggcatgc acttacactt caaacatcct tttgaaagac tgacaagttt 11640
gggtcttcac agttggaatt gggcatccct tttgtcaggg agggaggagc ggaggaggc 11700
tggtctgtta tgctgacaag tgtgattaaa ttcaaacttt gaggttaagt ggaggaactt 11760
gtacattgtt aggagtgtga caatttgac tcttaatgat ttggtcatac aaaatgaacc 11820
tagaccaact tctggaagat gtatataata actccatgtt acattgattt cacctgacta 11880
atacttatcc cttaacaatt aaatacagaa gatgccagcc atctgggctt ttaacccag 11940
aaatttagtt tcaaactcct aggttagtgt tctcactgag ctacatcctg atctagtcct 12000
gaaaatagga ccacatcac ccccaaaaa atctcaata agatttatgc tagtgtttca 12060
aaatttttagg aataggtaag attagaaagt ttttaatttt gagaaatggc ttctctagaa 12120
agatgtacat agtgaacact gaatggctcc taaagagcct agaaaactgg tactgagcac 12180
acaggactga gaggtcttcc ttgaaaagca tgtattgctt tacgtgggtc acagaaggca 12240
ggcaggaaga acttgggctg aaactggtgt cttaaagtgg taacatcttc acaactgat 12300
agcaagaact ttatcctgat gcaaaaacca ctaagtgaat ggtggcaatg 12360
gatccaggc tgctctagag gaggacttga cttctcatcc catcaccac accagatagc 12420
tcatagactg ccaattaaca ccagcttcta gcctccacag gcacctgcac tggtagacat 12480

```

aatttcacac aaacacagta agaagccttc cacctggcat ggtattgctt atcttttagtt 12540
cccaacactt gggaggcaga ggccagccag ggctatgtga caaaaacctt gtctagagga 12600
gaaacttcat agcttatttc ctattcacgt aaccagggtta gcaaaattta ccagccagag 12660
atgaagctaa cagtgtccac tatatttgta gtgttttaag tcaatttttt aaatatactt 12720
aatagaatta aagctatggg gaaccaagta caaacctggg gtattaactt gagaacttag 12780
cataaaaagt agttcatttg ttcagtaaat attaaatgct tactggcaaa gattatgtca 12840
ggaaacttgg aaatggatg gaaacaatca tagttgtaca tcttggttct gtgatcacct 12900
tggtttgagg taaaagtggg tcctttgatc aaggatggaa ttttaagttt atattcaatc 12960
aataatgtat tttttgtga ttgcaaaatt gcctatctag ggtataaaac ctttaaaaaat 13020
ttcataatac cagttcattc tccagttact aattccaaaa agccactgac tatggtgcca 13080
atgtggattc tgtttcctaa ggaaggattg tctgtgccct ttatttctaat agaacaatca 13140
cactgaaaaa ctaagctgaa agaagccaga ctttccctaaa taaataactt tccataaagc 13200
tcaaacaagg attactttta ggaggcactg ttaaggaact gataagtaat gaggttactt 13260
atataatgat agtcccacaa gactatctga ggaaaaatca gtacaactcg aaaaacagaac 13320
aaccagctag gcaggaataa cagggctccc aagtcaggag gtctatccaa cacccttttc 13380
tggtgagggc ccagaccta catattgtat acaaacaggg aggtgggtga ttttaactct 13440
cctgaggtag cttgtgctg agtaagcagt acagtgtaca gtttacattt 13500
tcatttaaaag atacattagc tccctctacc ccctaagact gacaggcact ttgggggtgg 13560
ggagggtctt ggaaaaatac gcttccatac actaaaagag aaatttcttt aattaggctt 13620
gttggttcca tacatctact ggtgtttcta ctacttagta atattataat agtcacacaa 13680
gcatctttgc tctgtttagg ttgtatatatt attttaaggc agatgataaa actgtagatc 13740
ttaagggatg cttctgcttc tgagatgata caaagaattt agaccataaa acagttaggt 13800
gcacaagcaa ctaaatatgg cctaaagtgt tctgacactt agaagccaag cagtgtaggc 13860
ttcttaagaa ataccattac aatcaccttg ctagaaatca agcattctgg agtggtaag 13920
cagtgttaacc tgtactgtaa gttacttttc tgctattttt ctcccaaaagc aagttcttta 13980
tgctgatatt tccagtgtta ggaactacaa atattaataa gttgtcttca ctcttttctt 14040
taccaaggag ggtctcttcc ttcacttga agtaaggat gaacaaaggc ttgagcagtg 14100
cgcttttagaa cataaactgc agcatgaagg ccccgatgt tcacccagac tacatggacc 14160
tttgcgcaca catgtcccat tccagataag gcctggcaca cacaaaaaac ataagtcatt 14220
aggctaccag tctgattcta aaacaaccta aaatcttccc acttaaatgc tatgggtggg 14280
gggttggaag gttgactcag aaaaactact gctgttttta gagaggatct ggggtcagtt 14340
tctgatacat tgtggcttac aactataact ccagttctag ggggtccatc caacatctc 14400
ttctgttagg ggcaccaa atgtgtattg tgtacaaaca gggagggtgag tgatttaact 14460
ctcgtgtata gtaccttggg aaaacatttc ttgtcctgag taagcagtag agctctgcct 14520
gtccctgggc tacagacacg gctcatttcc cgaaggcaag ctggatagag attccaattt 14580
ctcttcttgg atcccatcct ataaaagaag gtcaagttta atctattgca aaaggtaaat 14640
aggtagtttc ttacatgaga caagaacaaa tcttaggtgt gaagcagtag tcttttacag 14700
gccagagcct ctattctatg ccaatgaagg aaactgttag tccagtgtta tagagttagt 14760
ccagtgtata cttttctatc agaacacttt ttttttaaac aactgcaact tagcttattg 14820
aagacaaacc acgagtagaa atctgtccaa gaagcaagtg cttctcagcc tacaatgtgg 14880
aataggacca tgtaatggta cagtgaagtg aatgaattat ggcatgtttt tctgactgag 14940
aagacagtag aataaaagggt aaactcatgg tattttatta aaaagaatcc aatttctacc 15000
tttttccaaa tggcatatct gttacaataa tttccacaga agcagttctc agtgggaggt 15060
tgagatatac ccactgaaca gcatcaatgg gcaaacccca ggttggtttt ctgtggagac 15120
aaaggttaaga tttttcaata ttttttccca agctaagtag atggctcagc aaataatggg 15180
actggccatt aagtctcatg acctgagctt gatcctcagg gaccatgtgg tacaaggaga 15240
gacctaaatc cttcagttgg acttcaatct tctacctca tgtccacaca caaataaata 15300
caataaaaaa cattctgcag tcgaatttct aaaaggcgga at 15342

```

<210> 9

<211> 1823

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PGK-neo-pA

<400> 9

```

ggtcaattct accgggtagg ggaggcgctt ttcccaaggc agtctggagc atgcgcttta 60
gcagccccgc tgggcaactg gcgctacaca agtggcctct ggcctcgac acattccaca 120
tccaccggtg ggcgccaacc ggctccgttc tttgggtggc ccttcgcgc accttctact 180
cctcccctag tcagggaagt cccccccgcc ccgcagctcg cgctgtgag gacgtgacaa 240
atggaagtag cacgtctcac tagtctcgtg cagatggaca gcaccgctga gcaatggaag 300
gggttagggc tttggggcag cggccaatag cagctttgct ccttcgcttt ctgggctcag 360
aggctgggaa ggggtgggta cgggggcggt ctcaggcggt gctcaggggc gggggcggtg 420
cccgaaggta ctcgggagga ccggcattct gcacgcttca aaagcgcacg tctgccgcgc 480

```

```

tggtctcctc ttctcatct ccgggccttt cgacctgcag ccaatatggg atcgccatt 540
gaacaagatg gattgcacgc aggttctccg gccgcttggg tggagaggct attcgctat 600
gactgggcac aacagacaat cggctgctct gatgccgcg tgttccggct gtcagcgag 660
gggcgcccgg ttctttttgt caagaccgac ctgtccgggtg cctgaatga actgcaggac 720
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac 780
gttgctactg aagcgggaag ggactggctg ctattggcg aagtgccggg gcaggatctc 840
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg 900
ctgcatacgc ttgatccggc tacctgcccc ttcgaccacc aagcgaaaca tcgcatcgag 960
cgagcacgta ctcgatgga agcgggtctt gtcgacagg atgatctgga cgaagagcat 1020
caggggctcg cgccagccga actgttcgcc aggtcaagg cgcgcagcc cgacggcgag 1080
gatctcgtcg tgacctatgg cgatgcctgc ttgccgaata tcatggtgga aatggcgcg 1140
ttttctggat tcatcgactg tgccggctg ggtgtggcg accgctatca ggacatagcg 1200
ttggctaccc gtgatattgc tgaagagctt ggcgcgaaat gggctgaccg cttcctcgtg 1260
ctttacggta tcgcccgtcc cgattcgag cgcatcgct tctatcgct tcttgacgag 1320
ttcttctgag gggatcgatc cgctgtaagt ctgcagaaat tgatgatcta ttaaacaata 1380
aagatgtcca ctaaaatgga agtttttctt gtcatacttt gttaaagaag gtgagaacag 1440
agtacctaca ttttgatgg aaggattgga gctacggggg tgggggtggg gtgggattag 1500
ataaatgcct gctctttact gaaggtctt tactattgct ttatgataat gtttcatagt 1560
tgatatcat aatttaaaca agcaaaacca aattaagggc cagctcatc cccactca 1620
tgatctatag atctatagat ctctcgtagg atcattgtt ttctcttgat tccactttg 1680
tggttctaag tactgtggtt tccaaatgtg tcagtttcat agcctgaaga acgagatcag 1740
cagcctctgt tccacataca cttcattctc agtattgtt tgccaagttc taattccatc 1800
agaagctgac tctagatcct gca 1823

```

<210> 10

<211> 1867

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: CAGGS-Intron

<400> 10

```

tcgacattga ttattgacta gttattaata gtaatcaatt acggggtcat tagttcatag 60
cccataatag gatttccgag ttacataact tacggtaaat ggcccgcctg gctgaccgcc 120
caacgaccccc cgccattga cgtcaataat gacgtatgtt cccatagtaa cgccaatagg 180
gactttccat tgacgtcaat ggggtgacta tttacggtaa actgccact tggcagtaga 240
tcaagtgtat catatgcaaa gtacgcccc tattgacgtc aatgacggta aatggccgcg 300
ctggcattat gccagtaga tgacctatg ggactttcct acttggcagt acatctacgt 360
attagtcata gctattacca tgggtcgagg tgagccccc gttctgcttc actctcccca 420
tctccccccc ctcccccccc ccaattttgt atttatttat tttttaatta ttttgtgcag 480
cgatgggggg gggggggggg ggggcgcgcg ccaggcgggg cggggcgggg cgaggggcgg 540
tgccggggcg ggcggagagg tgccggcgca gccaatcaga gcggcgcgct ccgaaagtgt 600
ccttttatgg cgaggcggcg gcggcggcgg ccctataaaa agcgaagcgc gcggcgggcg 660
ggagtgcgtg cgttgccctt gccccgtgcc ccgctccgcg ccgcctcgcg ccgcccgcgc 720
cggctctgac tgaccgcgtt actcccacag gtgagcgggc gggacggccc ttctctccg 780
ggctgtaatt agcgtttggt ttaatgacgg ctcttttctt ttctgtggct gcgtgaaagc 840
cttaaagggc tccgggaggg ccctttgtgc gggggggagc ggctcggggg gtgcgtgcgt 900
gtgtgtgtgc gtggggagcg ccgcgtgcgg ccgcgcgtgc ccggcggctg tgagcgtgcg 960
gggcgcggcg cggggctttg tgccgtccgc gtgtgcgcga ggggagcgcg gccggggggc 1020
gtgcccgcgc gtgcgggggg gctgcgaggg gaacaaaggc tgctgtcggg gtgtgtgcgt 1080
gggggggtga gcagggggtg tgggcgcggc ggtcgggctg taaccccccc ctgcaccccc 1140
ctccccgagt tgctgagcac ggcccggctt cgggtgcggg gctcgtgcgc gggcgtggcg 1200
cggggctcgc cgtgccgggc ggggggtggc ggaggtggg ggtgccgggc ggggcggggc 1260
cgccctcggc cggggagggc tccggggagg ggcgcggcgg ccccgagcgc ccggcggtg 1320
tcgaggcgcg gcgagccgca gccattgcct tttatggtaa tcgtgcgaga gggcgagggg 1380
acttctcttg tcccaaatct ggcgagcggc aaatctggga ggccgcggcg caccctctct 1440
agcgggcgcg ggcaagcgg tgccgcggcg gcaggaagga aatggcgggg gagggccttc 1500
gtgcgtcgcc gdcgcggcgt ccccttctcc atctccagcc tcggggctgc cgcaggggga 1560
cggctgcctt cgggggggac ggggcagggc ggggttcggc ttctggcgtg tgaccggcgg 1620
ctctagtaag cgttgggggt agtactccct ctcaaaagcg ggcagactt ctgcgctaag 1680
attgtcagtt tccaaaaacg aggaggattt gatattcacc tggcccgcgg tgatgccttt 1740
gagggtggcc gcgtccatct ggtcagaaaa gacaatcttt ttgttgtcaa gcttgaggtg 1800
tggcaggctt gagatctggc catacacttg agtgacattg acatccactt tgcctttctc 1860
tccacag 1867

```

19/21

<210> 11
<211> 703
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: ZsGreen ORF

<400> 11
cgccaccatg gccaggtcca agcacggcct gaccaaggag atgaccatga agtaccgcat 60
ggagggtctgc gtggacggcc acaagttcgt gatcacggc gagggcatcg gctaccctt 120
caagggcaag caggccatca acctgtgctg ggtggagggc ggccccttgc ccttcgccga 180
ggacatcttg tccgcgcct tcatgtacgg caaccgcgtg ttcaccgagt acccccagga 240
catcgctcgac tacttcaaga actcctgccc cgccggctac acctgggacc gtccttcct 300
gttcgaggac ggccgcgtgt gcatctgcaa cgccgacatc accgtgagcg tggaggagaa 360
ctgcatgtac caccagttcca agttctacgg cgtgaacttc cccgccgacg gccccgtgat 420
gaagaagatg accgacaact gggagccctc ctgcgagaag atcatccccg tgccaagca 480
gggcatcttg aaggggcgacg tgagcatgta cctgctgctg aaggacggtg gccgcttgcg 540
ctgccagttc gacaccgtgt acaaggccaa gtccgtgccc cgcaagatgc ccgactggca 600
cttcatccag cacaagctga cccgcgagga ccgcagcgac gccagaacc agaagtggca 660
cctgaccgag caccgcatcg cctccggctc cgcttgccc tga 703

<210> 12
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: pA

<400> 12
aaataaagca atagcatcac aaatttcaca aataaa 36

<210> 13
<211> 7078
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: vector
consisting of a FRT-flanked neomycin resistance
gene and a PGK-TK-pA expression cassette

<400> 13
taaagggaac aaaagctgtc gagatctaga tatcgatggc catagagtta cgctagggat 60
aacagggttaa tatagcgcgc gtgccttagg taaccgaagt tcctatactt tctagagaat 120
aggaacttcg gaataggaac ttcttaggtc aattctaccg ggtaggggag gcgcttttcc 180
caaggcagtc tggagcatgc gcttttagcag ccccgctggg cacttggcgc tacacaagt 240
gcctctggcc tcgcacacat tccacatcca ccggtaggcg ccaaccggct ccgttctttg 300
gtggccctt cgccccaact tctactctc ccctagttag gaagttcccc ccgcgccgc 360
agctcgcgtc gtgcaggacg tgacaaatgg aagtagcacg tctcactagt ctgctgcaga 420
tggacagcac cgctgagcaa tgggaagcgg taggcctttg gggcagcggc caatagcagc 480
tttgcctctt cgttttctgg gctcagaggc tgggaagggg tgggtccggg ggcgggctca 540
ggggcgggct cagggcgggg gcggcgcccc gaaggtcctc cggaggcccc gcattctgca 600
cgcttcaaaa gcgcacgtct gccgcgctgt tctctcttc ctcatctccg ggcctttcga 660
cctgcagcca atatgggac ggccattgaa caagatggat tgcacgcagg ttctccggcc 720
gcttgggtgg agaggctatt cggctatgac tgggcacaac agacaatcgg ctgctctgat 780
gccgcgctgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg 840
tccggtgccc tgaatgaact gcaggacgag gcagcgcggc tatcgtggct gccacgacg 900
ggcgttctct gcgcagctgt gctcagactt gtcactgaag cgggaaggga ctggctgcta 960
ttgggcgaag tgccggggca ggatctctct tcacttcacc ttgctcctgc cgagaaagta 1020
tccatcatgg ctgatgcaat gcggcggtg catacgcttg atccggctac ctgcccattc 1080
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc 1140
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg 1200

ctcaaggcgc	gcatgcccga	cggcgaggat	ctcgtcgtga	cccatggcga	tgccctgcttg	1260
ccgaatatca	tggtggaaaa	tgcccgcttt	tctggattca	tcgactgtgg	ccggctgggt	1320
gtggcgacc	gctatcagga	catagcgttg	gctaccctg	atattgctga	agagcttggc	1380
ggcgaatggg	ctgaccgctt	cctcgtgctt	tacggtatcg	ccgctcccga	ttcgacgcgc	1440
atcgccctct	atcgccctct	tgacgagttc	ttctgagggg	atcgatccgc	tgtaagtctg	1500
cagaaaattga	tgatctatta	aacaataaag	atgtccacta	aaatggaagt	ttttcctgtc	1560
atactttgtt	aagaagggtg	agaacagagt	acctacattt	tgaatggaag	gattggagct	1620
acgggggtgg	gggtgggggtg	ggattagata	aatgcctgct	ctttactgaa	ggctctttac	1680
tattgcttta	tgataatggt	tcatagttgg	atatcataat	ttaaacaagc	aaaaccaa	1740
taagggccag	ctcattcctc	ccactcatga	tctatagatc	tatagatctc	tcgtgggagc	1800
attgtttttc	tcttgattcc	cactttgtgg	ttctaagtac	tgtggtttcc	aaatgtgtca	1860
gtttcatagc	ctgaagaacg	agatcagcag	cctctgttcc	acatacactt	cattctcagt	1920
attgttttgc	caagttctaa	ttccatcaga	agctgactct	agatcccgcg	ccgaagtacc	1980
tatactttct	agagaatagg	aacttcggaa	taggaacttc	aagcttaagc	gctagcacgt	2040
gttaactcga	gctgcaggat	tcgaggggcc	cggcaggtca	attctaccgg	gtaggggagg	2100
cgcttttccc	aaggcagctc	ggagcatgog	ctttagcagc	cccgctgggc	acttggcgct	2160
acacaagtgg	cctctggcct	cgcacacatt	ccacatccac	cggtaggcgc	caaccggctc	2220
cgttctttgg	tgccccttc	gcgccacctt	ctactcctcc	cctagtccag	aagttcccc	2280
ccgccccgca	gctcgcgctg	tgacaggagt	gacaaatgga	agtagcacgt	ctcactagtc	2340
tcgtgcagat	ggacagcacc	gctgagcaat	ggaagcgggt	aggcctttgg	ggcagcggcc	2400
aatagcagct	ttgtccttc	gctttctggg	ctcagaggct	gggaagggtg	gggtccgggt	2460
gcgggctcag	ggcggggctc	agggcgggg	cgggcgcccg	aaggtcctcc	ggaggccggg	2520
cattctgcac	gcttcaaaag	cgacgtctg	ccgcgtgtt	ctcctcttcc	tcattctcgg	2580
gcctttcgac	ctgcagccaa	tgacccgtcc	ttgccatcat	ggcctcgtac	cccggccatc	2640
aacacgcgtc	tgcgttcgac	caggctgcgc	gttctcgcgg	ccatagcaac	cgacgtacgg	2700
cgttgccgcc	tcgcccggcag	caagaagcca	cggaaagtcg	cccggagcag	aaaatgcccc	2760
cgctactcgc	ggtttatata	gacggtcccc	acgggatggg	gaaaaccacc	accacgcaac	2820
tgctggtggc	cctgggttcg	cgcgacgata	tcgtctacgt	acccgagccg	atgacttact	2880
ggcggtgct	gggggcttcc	gagacaatcg	cgaacatcta	caccacacaa	caccgcctcg	2940
accagggtga	gatatcgggc	ggggacgcgg	cggtggtaat	gacaagcgcc	cagataacaa	3000
tgggcatgcc	ttatgccgtg	accgacgcgg	ttctggctcc	tcatatcggg	ggggaggctg	3060
ggagctcaca	tgcccgcgcc	ccggccctca	cctcatctt	cgaccgcgat	cccatcgccg	3120
ccctcctgtg	ctaccgggcc	gcgcggtacc	ttatgggcag	catgaccccc	caggccgtgc	3180
tggcgttcgt	ggccctcatc	ccgcccagct	tgcccggcac	caacatcgtg	cttggggccc	3240
ttccggagga	cagacacatc	gaccgcctgg	ccaaacgcca	gcgcccgggc	gagcggctgg	3300
acctggctat	gctggctgcg	attcgccgcg	tttacgggct	acttgccaat	acgggtgcgtg	3360
atctgcagtg	cgccgggtcg	tgccggggag	actggggaca	gctttcgggg	acggccgtgc	3420
cgcccccagg	tgccgagccc	cagagcaacg	cgggcccaag	accccatatc	ggggacacgt	3480
tatttaccct	gtttcgggcc	cccaggttgc	tgccccccaa	cggcgacctg	tataacgtgt	3540
ttgcctgggc	cttggaacgtc	ttggccaaac	gcctccgttc	catgcacgtc	tttatcctgg	3600
attacgacca	atcgcccggc	ggctgcgggg	acgccttgc	gcaacttacc	tcggggatgg	3660
tcgacacca	cgtaaccacc	cccggctcca	taccgaagat	atgcgacctg	gcgcgcacgt	3720
ttgcccggga	gatgggggag	gctaactgag	gggatcgatc	cgctcgttaa	gtctgcagaa	3780
attgatgatc	tattaaacaa	taaagatgtc	cactaaaatg	gaagtttttc	ctgtcatact	3840
ttgttaagaa	gggtgagaac	agagtacctt	cattttgaat	ggaaggattg	gagctacggg	3900
ggtgggggtg	gggtgggatt	agataaatgc	ctgctcttta	ctgaaggctc	tttactattg	3960
ctttatgata	atgtttcata	gttgatatac	ataattttaa	caagcaaaac	caaatttaag	4020
gocagctcat	tcctccact	catgatctat	agatctatag	atctctcgtg	ggatcattgt	4080
ttttctcttg	attcccactt	tgtgggttcta	agtactgtgg	tttccaaatg	tgtcagtttc	4140
atagcctgaa	gaacgagatc	agcagcctct	gttccacata	cacttcattc	tcagtattgt	4200
tttgccaagt	tctaattcca	tcagaagctg	actctaggcc	gagctccaat	tcgcccata	4260
gtgagtcgta	ttacaattca	ctggccgctg	ttttacaacg	tcgtgactgg	gaaaaccctg	4320
gogttaccca	acttaatcgc	cttgacgac	atcccccttt	cgccagctgg	cgtaaatagc	4380
aagaggcccg	caccgatcgc	ccttcccaac	agttgcgcag	cctgaatggc	gaatgggacg	4440
cgccctgtag	cggcgcatta	agcgcggcgg	gtgtgggtgg	tacgcgcagc	gtgaccgcta	4500
cacttgccag	cgccctagcg	cccgtctcct	tcgctttctt	cccttctctt	ctcgccacgt	4560
tcgcccggct	tccccgtcaa	gctctaaatc	gggggctccc	tttaggggtc	cgatttagtg	4620
ctttacggca	cctcgacccc	aaaaaacttg	attagggtga	tggttcacgt	agtgggccat	4680
cgccctgata	gacgggtttt	cgccctttga	cgttgagtc	cacgttcttt	aatagtgagc	4740
tcttggtcca	aactggaaca	acactcaacc	ctatctcggt	ctattctttt	gatttataag	4800
ggattttgcc	gatttcggcc	tattgggttaa	aaaatgagct	gatttaacaa	aaatttaacg	4860
cgaatttttaa	caaaatatta	acgcttacaa	tttaggtggc	acttttcggg	gaaatgtgcg	4920
cggaaaccct	atttgtttat	ttttctaaat	acattcaaat	atgtatccgc	tcattgagaca	4980
ataaccctga	taaatgcttc	aataatattg	aaaaaggaag	agtatgagta	ttcaacattt	5040
ccgtgtcgcc	cttattccct	tttttgccg	cctgtttttg	ctcaccacga	gttacatcga	5100
aacgctggtg	aaagtaaaag	atgctgaaga	tcagttgggt	gcacgagtg	gttttccaat	5160
actggatctc	aacagcggtg	agatccttga	gagttttcgc	ccggaagaac		5220

gatgagcact	tttaaagttc	tgctatgtgg	cgcggtatta	tcccgtattg	acgccgggca	5280
agagcaactc	ggtcgccgca	tacactattc	tcagaatgac	ttggttgagt	actcaccagt	5340
cacagaaaag	catcttacgg	atggcatgac	agtaagagaa	ttatgcagtg	ctgccataac	5400
catgagtgat	aacactgcgg	ccaacttact	tctgacaacg	atcggaggac	cgaaggagct	5460
aaccgctttt	ttgcacaaca	tgggggatca	tgtaactcgc	cttgatcggt	gggaaccgga	5520
gctgaatgaa	gccataccaa	acgacgagcg	tgacaccacg	atgcctgtag	caatggcaac	5580
aacgttgcg	aaactattaa	ctggcgaaact	acttactcta	gcttcccggc	aacaattaat	5640
agactggatg	gagggcgata	aagttgcagg	accacttctg	cgctcggccc	ttccggctgg	5700
ctggtttatt	gctgataaat	ctggagccgg	tgagcgtggg	tctcgcggta	tcatgtcagc	5760
actggggcca	gatggtaagc	cctcccgtat	cgtagttatc	tacacgacgg	ggagtcaggc	5820
aactatggat	gaacgaaata	gacagatcgc	tgagataggt	gcctcactga	ttaagcattg	5880
gtaactgtca	gaccaagttt	actcatatat	actttagatt	gatttaaaac	ttcattttta	5940
attttaaagg	atctaggtga	agatcctttt	tgataatctc	atgacaaaaa	tcccttaacg	6000
tgagttttcg	ttccactgag	cgtcagaccc	cgtagaaaaa	atcaaaggat	cttcttgaga	6060
tccttttttt	ctgcgcgtaa	tctgctgctt	gcaaacaaaa	aaaccaccgc	taccagcggg	6120
ggtttggttg	coggatcaag	agctaccaac	tctttttccg	aaggttaactg	gcttcagcag	6180
agcgagata	ccaaataactg	tccttctagt	gtagccgtag	ttaggccacc	acttcaagaa	6240
ctctgtagca	ccgcctacat	acctcgctct	gctaatacctg	ttaccagtgg	ctgctgccag	6300
tggcgataag	togtgtctta	ccgggttgga	ctcaagacga	tagttaaccg	ataaggcgca	6360
gcggtcgggc	tgaacggggg	gttcgtgcac	acagcccagc	ttggagcgaa	cgacctacac	6420
cgaactgaga	tacctacagc	gtgagctatg	agaaagcgcc	acgcttcccg	aaggagagaa	6480
ggcggaacag	tatccggtaa	gcggcagggt	cggaacagga	gagcgacaga	gggagcttcc	6540
agggggaaac	gcctggatc	tttatagtcc	tgctgggttt	cgccacctct	gacttgagcg	6600
togatttttg	tgatgctcgt	cagggggggc	gagcctatgg	aaaaacgcca	gcaacgcggc	6660
ctttttacgg	ttcctggcct	tttgctggcc	ttttgctcac	atgttctttc	ctgcgttatc	6720
ccctgattct	gtggataacc	gtattaccgc	ctttgagtga	gctgataccg	ctcgccgca	6780
ccgaacgacc	gagcgacgcg	agtcagttag	cgagggaagc	gaagagcgcc	caatacgcaa	6840
accgcctctc	ccgcgcggtt	ggccgattca	ttaatgcagc	tggcacgaca	ggtttcccga	6900
ctggaaagcg	ggcagtgagc	gcaacgcaat	taatgtgagt	tagctcactc	attaggcacc	6960
ccaggcttta	cactttatgc	ttccggctcg	tatgttggtg	ggaattgtga	gcggataaca	7020
atttcacaca	ggaaacagct	atgaccatga	ttacgccaa	cgcgcaatta	accctcac	7078

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/EP2004/000968

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/82 C12N15/90 C12N15/65

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>US 5 487 992 A (THOMAS KIRK R ET AL) 30 January 1996 (1996-01-30) cited in the application abstract figure 7C column 2, paragraph 4 -column 3, paragraph 2 column 5, paragraph 2 - paragraph 3 column 6, paragraph 2 column 7, paragraph 2 - paragraph 3 column 10, paragraph 2 - paragraph 3 column 11, paragraph 2 - paragraph 3 column 13, paragraph 3 column 15 column 16, paragraph 2 - paragraph 3 column 17, paragraph 3 example 5</p> <p style="text-align: center;">----- -/--</p>	1-12

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *Z* document member of the same patent family

Date of the actual completion of the international search

11 May 2004

Date of mailing of the international search report

01/06/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Keller, Y

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP2004/000968

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 01/66717 A (RONG YIKANG S ;UNIV UTAH (US); DREWS GARY N (US); GOLIC KENT G (US) 13 September 2001 (2001-09-13) cited in the application abstract figure 10 page 12, paragraph 1 ----	1-12
X	WO 01/95711 A (DELTAGEN INC ;BRENNAN THOMAS J (US); GUENTHER CATHERINE (US)) 20 December 2001 (2001-12-20) page 8, paragraph 3 -page 10, paragraph 3 claims 1,2 -----	1-12

INTERNATIONAL SEARCH REPORT

 International Application No
 PCT/EP2004/000968

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5487992	A	30-01-1996	US 5464764 A	07-11-1995
			US 5627059 A	06-05-1997
			US 6689610 B1	10-02-2004
			US 6204061 B1	20-03-2001
			US 5631153 A	20-05-1997
<hr/>				
WO 0166717	A	13-09-2001	AU 4728501 A	17-09-2001
			CA 2401677 A1	13-09-2001
			EP 1259600 A2	27-11-2002
			WO 0166717 A2	13-09-2001
<hr/>				
WO 0195711	A	20-12-2001	AU 6693601 A	24-12-2001
			CA 2410659 A1	20-12-2001
			EP 1294227 A2	26-03-2003
			JP 2004503226 T	05-02-2004
			WO 0195711 A2	20-12-2001
			US 2002056146 A1	09-05-2002

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☒ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.